U.PORTO

FEP FACULDADE DE ECONOMIA UNIVERSIDADE DO PORTO

# Skill Mismatches and Labor Market Outcomes

Maria Isabel Da Silva Araújo

January 2019

Doctoral Thesis in Economics

Supervisor: Anabela de Jesus Moreira Carneiro

## Biography

Maria Isabel Da Silva Araújo was born on January 26, 1973. She lived in her hometown in Paris (France) until 1998, after which she moved to Portugal. She concluded her undergraduate studies in International Relations in 2002, at the University of Minho. During her undergraduate studies she received two merit scholarships from University of Minho for her academic achievement. She joined the Master of Science (MSc) in Monetary, Banking and Financial Economics at that same institution in 2011. Her MSc dissertation "Gold as a Safe Haven" was supervised by Professors Cristina Amado and Fernando Alexandre. In 2014, she enrolled the Doctoral Programme at the Faculty of Economics of University of Porto (FEP), under the supervision of Professor Anabela de Jesus Moreira Carneiro.

From 2002 to 2009 she worked as controller and purchasing manager in a textile Company. In 2013 she worked as Financial Manager at ITIM Group – Profimetrics Software Solutions SA.

In 2010 she taught Economics at Escola Secundária de Barcelinhos. In 2011 she joined the Polytechnic Institute of Cávado and Ave (IPCA) as a lecturer, teaching the courses of Economics, Financial Calculation and Financial Management. In 2016 she also joined the Polytechnic of Porto (ISCAP) as a lecturer, teaching the courses of Economics and Intercultural Economic Relations.

## Acknowledgements

After four long years of challenging and intensive studies in economics, I would like to thank all my professors of the School of Economics and Management at the University of Porto. A special note of gratitude is devoted to Professor Elvira Silva for her guidance and support during this journey. I feel extremely honoured for being her student and for her friendship.

I have no words to express my gratitude to my supervisor Anabela Carneiro for the countless hours spent in Quadros de Pessoal, sharing her experience and knowledge with me. As well as mentoring and supporting me through this endeavour endlessly. I will be eternally greatful to her for guiding me into becoming a researcher, it has been such an honour to work and learn from her. I am terribly fortunate for having such an outstanding person as supervisor.

I have also acknowledged all the suggestions that were provided during my presentations in workshops and conferences during my Ph.D course. A special thanks goes to Professor Miguel Portela for reviewing and helping me improve this work. I am grateful to my fellow colleagues at IPCA and ISCAP for their motivation.

I would like to thank my dear friends for their support, love and understanding: I'm back! I would also like to acknowledge all the help and motivation I received from my friends of the Ph.D in Economics at FEP, in particular Ana Isabel, Claúdia, Filipa, Sandra, Francisco, André, António, José Gaspar. It has been a long road and we did it together. I will miss you and keep you in my heart.

Last but not least, I would like to thank my family, husband Pedro, children Sofia and Tomás, to my parents, brother, nieces, cousins, sisters-in-law, brother-in-law and parents-in-law, for all the love and support I was given during this journey. This work would not have been possible without you. I love you all to infinity and beyond.

## Abstract

This dissertation comprises three essays on skill mismatches.

The first essay studies the incidence of the over- and undereducation phenomenon and its impact on individual wages using longitudinal data for Portugal over the 1995-2012 period and different measures of educational mismatches. In this essay, we document the incidence of over- and undereducation in Portugal in the 1995-2012 period and its recent time trends using different measures of educational mismatches. We also exploit a novel measure of required schooling based on realized matches for the flows of new hires. Finally we take into account the role of worker and firm unobserved heterogeneity in our estimations of the wage effects of educational mismatches. We found that half of the employees in the private sector in Portugal experienced a vertical educational mismatch in the analysed period. This proportion increases to almost three fourths when an exogenous measure based on the international standard classification of occupations and education is used. We also found that accounting simultaneously for worker and firm permanent observed and unobserved heterogeneity reduces dramatically the returns to over- and underducation, suggesting that educational mismatches are largely driven by unobserved characteristics of the worker and the firm and failure to account for them bias the estimates of the mismatch educational effects. Finally we found that undereducated seem to correspond to a higher ability group of employees, while the overeducated seem to correspond to a low-ability group of workers.

In the second essay, using a sample of higher education graduates in Portugal, we analyse the career dynamics of workers who entered the labor market for the first time in a job for which they were overqualified. We identify the transitions out of overqualification of 13,709 recent graduates who entered the labor market overqualified in 2006 or 2007 and analyse the determinants of the transitions out of overqualification considering the duration of the first spell in a mismatched job. Our measure of overqualification is based on the relative importance of tasks within each occupation at a 2 digit-level according to the O*NET classification. Then we examine whether being overqualified in the first job leaves a scar on future wages of recent graduates (controlling for both workers individual observed and unobserved permanent heterogeneity and firm observed and unobserved permanent heterogeneity). We found that overqualification is a permanent phenomenon for a great majority of workers: six years after entering the labor market, 63% of the workers that entered overqualified remain in that status. Finally we found that at entry, overqualified workers earn lower wages when compared with well-matched workers, but this gap tends to diminish for overqualified workers who were able to move to a well-matched job. In fact, we found that overqualified workers that switched to a well-matched job experience a wage growth that exceeds the wage growth of their similar well-matched counterparts in 12 percentage points, once controlling for workers and firms observed and unobserved permanent heterogeneity.

Finally, in the third essay, we propose to compare the mobility pattern of recent graduates who entered the labor market for the first time in 2006 or 2007 overqualified with the mobility pattern of recent graduates who entered the labor market for the first time in 2006 or 2007 well-matched. The data indicate that the early career of young graduate workers in Portugal is characterized by low job mobility. In fact, we found that more than half of the graduates who entered the labor market for the first time in 2006 or 2007 did not change their first employer over the analysed period. Using the number of employer changes as an indicator of job mobility, the econometric approach revealed that there is no statistical difference in terms of the total number of job changes experienced by overqualified workers at entry and well matched workers at entry. However, when we consider a finer measure of job mobility - the total number of job changes to a well-matched job - the estimates indicate that overqualified workers at entry are less likely to switch to a well-matched job.
**Keywords:** educational mismatches, skill mismatches, overeducation, overqualification, wages, persistence.

**JEL Classification Numbers:** I23; I26; J24; J31; J62; J63

# Resumo

Esta dissertação compreende três ensaios sobre os desajustamentos entre os níveis de qualificações dos trabalhadores portugueses e os níveis de qualificações exigidos nas ocupações e os seus efeitos sobre os salários e o emprego.

Usando os dados longitudinais dos Quadros de Pessoal para o período de 1995 a 2012, no primeiro ensaio estuda-se a incidência da sobre e sub-educação em Portugal assim como o seu impacto nos salários. Explorando diferentes medidas para medir os desajustamentos de escolaridade, os resultados indicaram que cerca de metade dos trabalhadores portugueses no setor privado experimentaram um desajustamento educacional na ocupação no período analisado. Os resultados revelaram ainda que os retornos salariais à sobre e sub-educação são substancialmente reduzidos quando se controla para as características não observáveis das empresas e dos trabalhadores, sugerindo que os desajustamentos educacionais estão correlacionados com as características não observáveis dos trabalhadores e das empresas.

Utilizando uma amostra de recém-diplomados do ensino superior, no segundo e terceiro ensaios analisam-se as dinâmicas de carreira ao longo do período de 2006 a 2012 de duas coortes de jovens diplomados que entraram pela primeira vez no mercado de trabalho em Portugal em 2006 ou 2007, distinguindo o seu estatuto à entrada no que respeita ao desajustamento entre as suas qualificações e as requeridas na respetiva ocupação. Nestes dois ensaios a medida de sobre-qualificação é baseada na importância relativa das tarefas dentro de cada ocupação a um nível de 2 dígitos, de acordo com a classificação da O*NET.

Em particular, no segundo ensaio, e com o objetivo de analisar a persistência do fenómeno da sobre-qualificação são identificadas as saídas da condição de sobre-qualificado de 13,709 recém-diplomados que entraram no mercado de trabalho sobre-qualificados, considerando a duração do primeiro emprego. A análise empírica revelou que, ao fim de seis anos, 63% dos indivíduos que entram no mercado de trabalho sobre-qualificados permanecem nesse estatuto. Finalmente, e no que respeita aos salários, os dados revelam salários mais baixos à entrada para os trabalhadores sobre-qualificados quando comparados com os salários dos trabalhadores que entraram no mercado de trabalho numa ocupação que exige os seus níveis de qualificação. No entanto, controlando para a heterogeneidade individual não observada das emnpresas e dos trabalhadores, as estimativas indicam que os trabalhadores sobre-qualificados que acabam por transitar para uma ocupação adequada às suas qualificações experimentam, em média, um crescimento salarial superior ao dos trabalhadores adequadamente qualificados à entrada reduzing o gap salarial inicial entre estes.

Finalmente, no terceiro ensaio o objetivo é comparar o padrão de mobilidade inter-

empresas dos recém-diplomados que entraram no mercado de trabalho pela primeira vez em 2006 ou 2007 sobre-qualificados na ocupação com o padrão de mobilidade dos recém-diplomados que entraram no mercado de trabalho pela primeira vez em 2006 ou 2007 adequadamente qualificados na ocupação. Os dados indicam que o início da carreira dos jovens licenciados em Portugal é caracterizado por uma baixa mobilidade profissional. De facto, verifica-se que mais de metade dos graduados não mudaram de empregador durante o período analisado. Usando o número de mudanças de empregador como um indicador da mobilidade profissional, a abordagem econométrica revelou que não há diferença estatística significativa em termos do número total de mudanças de emprego experimentadas por trabalhadores sobre-qualificados à entrada e trabalhadores similares adequadamente qualificados à entrada. No entanto, quando consideramos uma medida mais refinada da mobilidade profissional - o número total de mudanças de emprego para uma ocupação adequada às qualificações - as estimativas indicam que os trabalhadores sobre-qualificados à entrada são menos propensos a mudar para uma ocupação que se adeque às suas qualificações quando comparados com trabalhadores similares que entraram no mercado de trabalho adequadamente qualificados.
**Palavras-chave**: desajustamentos de escolaridade, desajustamento de qualificações, sobre-educação, sobre-qualificação, salários, persistência
**Código JEL**: I23; I26; J24; J31; J62; J63

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Educational Mismatches and Wages: Accounting for Worker and Firm Unobserved Heterogeneity

Abstract: Exploring a rich matched employer-employee data set over the 1995-2012 period, this study provides novel evidence on the effects of educational mismatches on wages. Using different criteria to measure over- and undereducation, the data show that more than 50 percent of the employed in the private sector in Portugal experienced a job mismatch. According to the job analysis approach, overeducation incidence exibhits an increasing trend, while undereducation incidence a decreasing one. According to the statistical measures based on the mode of the stock of employees and the mode of the flow of new hires, overeducation is decreasing and undereducation is increasing indicating that labor market demand is keeping pace with the rise in educational attainment of the portuguese population. Regarding wages, the results reveal that the wage differential between adequately matched workers and mismatched workers tends to vanish once worker and firm unobserved heterogeneity is taken into account. Furthermore, the data indicate that the overeducated are the least-able employees, while the undereducated the most-able, suggesting that educational mismatches are largely driven by individual unobserved attributes.

KEYWORDS: educational mismatches, overeducation, undereducation, wages, two-way fixed effects

JEL CODES: I26; J24; J31

## 1.1 Introduction

Skill mismatches in the labor market are a matter of great concern for academics, practitioners, and policymakers. Skill mismatches arise from several imbalances between the skills offered and the skills demanded in the labor market due to search and matching frictions such as information asymmetries, mobility costs, and unresponsive education and training systems to the world of work (Quintini, 2011; ILO, 2014). For firms, these imbalances lead to inefficiencies in the utilization of labor, with detrimental effects on productivity (Tsang and Levin, 1985) and turnover rates (Sicherman, 1991; Hersch, 1991). For workers, they can affect job satisfaction (Tsang, Rumberger, and Levin, 1991; Maynard, Joseph, and Maynard, 2006; Erdogan, Bauer, Peiró, and Truxillo, 2011), training investments (Groot, 1993; Van Smoorenburg and Van der Velden, 2000) and wages (Hartog, 2000; Groot and Van den Brink, 2000). Positive outcomes of overqualification associated with training and job performance are also documented in a few studies (e.g., Fine, 2007; Fine and Nevo, 2008).

ILO (2014) report showed that the level of skill mismatch is considerable in Europe (between 30 and 50 percent of the employed in the European countries are mismatched), with overeducation rising and undereducation decreasing in the majority of the countries studied.

Previous studies investigated the earnings consequences of over- and undereducation augmenting the standard mincerian wage equation with measures of over- and underschooling (Duncan and Hoffman, 1981; Rumberger, 1987; Verdugo and Verdugo, 1989; Sicherman, 1991; Cohn and Kahn, 1995 for the U.S; Hartog and Oosterbeek, 1988 for the Netherlands; Alba-Ramirez, 1993 and Murillo, Rahona-López, and Salinas-Jiménez, 2012 for Spain; Kiker, Santos, and Oliveira, 1997, and Oliveira, Santos, and Kiker, 2000 for Portugal; Cohn and Ng, 2000 for Hong Kong; Dolton and Vignoles, 2000 for the U.K.). This first wave of empirical studies, drawn on cross-section data, established two basic stylized facts. First, based on the Duncan and Hoffman (1981) model (ORU specification), the empirical results show that, when compared with their job co-workers who are adequately educated, overeducated workers receive a wage bonus for the extra years of surplus schooling and undereducated workers a wage penalty for the deficit years of schooling, even though smaller than the returns to required education. Second, based on the Verdugo and Verdugo (1989) model, the empirical evidence showed that overeducated workers earn less and undereducated workers earn more than their counterparts with the same years of schooling, but who hold jobs for which they are adequately educated.

Overall, these findings support the hypothesis that wages are not uniquely determined on the basis of the individual's educational level. The characteristics of the job seem also to play an important role in wage determination favoring an assignment theory interpretation of the labor market (Sattinger, 1993). According to the assignment theory, productivity depends on both - worker and job characteristics - and is maximised when workers are allocated to the jobs according to their skill ranks. Most skilled workers should be assigned to the most complex and demanding jobs, whereas the least skilled to the simplest jobs. In this setup, workers with the same attained level of education may perform differently depending on the job they hold, allowing individuals to make their choices according to their relative comparative advantage in the possible set of job offers.

In the last two decades, a second wave of studies, drawn on longitudinal data, further explored the wage effects of educational mismatches. Overall, this literature showed that the wage gap between well-matched and mismatched workers reduces significantly or even disappears, once unobserved individual heterogeneity is taken into account (see Bauer, 2002 for Germany; Frenette, 2004 for Canada; Tsai, 2010 for the U.S.; Mavromaras, McGuinness, O'Leary, Sloane, and Wei, 2013 for Australia), suggesting the existence of a trade-off between over- and undereducation and other components of human capital (Sicherman, 1991).

The purpose of this paper is to extend this recent literature, which emphasizes the role of worker unobserved heterogeneity, by incorporating in the analysis the role of firm observed and unobserved permanent heterogeneity. Recruitment and selection entail a two-sided matching process between the firm and the worker that are simultaneously determined by observed and unobserved characteristics of both parties. The reasons why an employer may decide to hire a worker that is apparently underqualified for a given job or why a worker may accept a job for which his/her attained level of education exceeds the level required for the job may be driven by unobserved characteristics of the worker and the employer. Failure to account for endogeneity in these choices may bias the estimates of educational mismatches on wages.

In this study we acknowledge at least two endogeneity problems that should be addressed in our empirical methodology. The first stems from the fact that individuals are heterogenous and differ in certain unobserved characteristics. If these unobserved characteristics are correlated with the mismatch status, neglecting them may bias the estimated effects of educational mismatches on earnings (Chevalier, 2003; McGuinness, 2006). Actually, some studies found evidence that ability and overschooling are negatively correlated (e.g., Greene and McIntosh, 2007; Chevalier and Lindley, 2009).

The second potential source of endogeneity may arise from selection effects since workers sorting across firms is non-random (Baptista, Lima, and Preto, 2013; Rocha, van Praag, Folta, and Carneiro, 2018, in press). In this framework, workers self-selection imply that individuals make choices regarding whether they join a given firm/job based on their own attributes and firm/job characteristics (most often not observed by the researcher) that may be correlated with both educational mismatches and expected wages. Suppose that workers are more likely to accept a job for which his/her level of education exceeds the required for the job in large/high-paying firms that offer more career prospects and a higher wage growth. In this case, ignoring firm heterogeneity may bias the estimates of the returns to overeducation as high-paying firms may be more likely to attract workers that are willing to accept a job for which they are overeducated. Employers selection is also a concern in this study. Recent studies on new ventures performance, have shown that hiring decisions are not independent from founders observed and unobserved characteristics (Rocha *et al.*, 2018, in press). Suppose that business-owners or managers less qualified, risk averse or with low levels of confidence may be more likely to recruit candidates who have excess education. Neglecting these issues, potentially correlated with job mismatches and expected outcomes, may mislead the causal interpretation of the effects of over- and undereducation on wages.

Exploring a rich matched employer-employee dataset, this paper revisits the literature on the wage effects of educational mismatches by incorporating in the analysis controls for firm observed and unobserved permanent heterogeneity. Thus, our contribution to this literature relies on two main ingredients: (i) a rich administrative longitudinal linked employer-employee data set, and (ii) an identification strategy that allows to simultaneously account for worker and firm unobserved heterogeneity.

We are confident that Portugal constitutes an interesting case to develop this exercise. First, Portugal experienced, in the past two decades, a huge increase in the educational levels of its population, favored by the massive expansion of the higher education system (Figueiredo, Teixeira, and Rubery, 2013). Nevertheless, and despite the significant improvement in the educational levels of the labor force, Portugal still ranks below the OECD and UE averages in terms of schooling attainment (OECD, 2017). Second, the portuguese labor market is characterized by very strict employment regulation legislation, where individual and collective dismissals of workers in permanent contracts are the most restrictive across the OECD (OECD, 2012; Martins, 2009;

Centeno and Novo, 2012). The termination of a permanent contract in Portugal involves a lengthy and complex administrative procedure that imposes several costs to firms and creates barriers against workforce adjustments (Martins, 2009) that potentially difficults match quality improvements. Third, Portugal seems to exibhit, since the mid 1990s, a phenomenon of job polarization in the labor market mainly driven by technological changes (Fonseca, Lima, Pereira, 2018) that enhances the recruitment of better-educated cohort of workers. Using *Quadros de Pessoal* data for 1986-2007, Fonseca *et al.* (2018) show a sharp increase of employment in abstract tasks relative to manual tasks, along with a decline for routine manual tasks.

In sum, the contribution of this study to the current literature is threefold. First, to document the incidence of over- and undereducation in Portugal in the 1995-2012 period and its recent time trends using different measures of educational mismatches. Second, to exploit a novel measure of required schooling based on realized matches for the flows of new hires. Third, to complement prior research that aimed to estimate the economic returns to over- and undereducation by taking simultaneously into account the role of worker and firm unobserved heterogeneity.

The rest of the paper is organized as follows. Section 1.2 describes the data and reports figures of the incidence of over- and undereducation in Portugal in the 1995-2012 period. Section 1.3 describes the econometric model and discusses identification issues. Empirical results and robustness checks are presented and discussed in Section 1.4. Section 1.5 concludes.

## 1.2 Data and Methodological Issues

### 1.2.1 Quadros de Pessoal (QP)

Our data comes from Quadros de Pessoal (*QP*), a matched employer-employee dataset collected by the Portuguese Ministry of Labor, Solidarity, and Social Security. *QP* is an annual mandatory survey that all Portuguese firms, in the private sector, with at least one wage earner are legally obliged to fill in.[1] Data are available from 1985 until 2012 and include information at the firm, establishment, and worker level. At the firm level,

[1] Public administration, self-employment and nonmarket services are not covered by *QP*.

*QP* contains information on industry, location, employment, sales, ownership, legal setting, among others. Worker data include information on gender, age, education, occupation, qualification, tenure, wages, hours worked, among others.

Firms, establishments, and workers entering the database are assigned a unique identifying number that makes it possible to track them across all annual waves of data. Furthermore, the worker files includes the firm and establishment number to which each individual is affiliated in a given year, allowing to match workers with their employers both at the firm and the establishment level.

The longitudinal nature of the dataset, the long time span covered and its high degree of representativeness and reliability, makes *QP* an appropriate database for a comprehensive study on the wage effects of educational mismatches.

In this study, we will use worker-level data from 1995 to 2012.[2] Our dataset includes the population of wage earners employed in mainland Portugal aged between 15 and 64 years old.[3] Observations with missing and inconsistent values in gender, age, and education were repaired whenever possible, or dropped otherwise.[4] The data include 27,268,443 (years$\times$individuals) observations for both genders, which correspond to 4,209,971 individuals and around 458,342 firms. The average number of periods per individual is 6.5 years.

### 1.2.2 Measuring Educational Mismatches

The measure used to define educational mismatch among the employed is crucial to our analysis and previous literature showed that the patterns of skill mismatches strongly depend on the criteria adopted to measure mismatches (e.g., ILO, 2014; Groot *et al.*, 2000). In this paper we will focus on vertical educational mismatches in the Portuguese labor market. A vertical mismatch takes place when the level of education is higher or lower than the one required for the job. Following previous studies, we will rely on statistical measures based on realized matches and on a measure based on job analysis

[2] Worker data are not available for the year 2001.

[3] Workers in agricultural and fishery were excluded from the analysis,as well as workers in the islands of Açores and Madeira due to the lack of representativeness.

[4] The top and bottom 1 percent observations in wages were excluded from the sample. Multiple job-holder workers were also dropped.

to identify the required level of education for a given job (e.g., Verdugo and Verdugo, 1989; Oliveira *et al.*, 2000; Hartog and Groeneveld, 2004; Korpi and Tåhlin, 2009; Bauer, 2002). To apply the statistical measures, first we need to convert the completed levels of schooling available in *QP* in a quantitative variable measured in years of schooling. The Portuguese educational system is structured in three levels: primary (first cycle, second cycle, third cycle), secondary (secondary and post-secondary), and tertiary education (bachelor, graduate, master, doctoral degree). We assigned years of schooling as follows: less than first cycle - 0 years; first cycle - 4 years; second cycle - 6 years; third cycle - 9 years; secondary education and post-secondary education - 12 years; bachelor - 15 years; graduate, master degree, doctoral degree - 16 years.[5]

Following Kiker *et al.* (1997), our first measure of required education based on realized matches is defined as the mode level of education for the stock of workers employed in a three-digit occupation in a given year.[6]

The second measure, a novelty in this study, takes more properly into account that the level of required education may change over the years as employers hiring standards adjust to technological changes, organizational structure changes, increases in the relative supply of higher educated workers, etc. In this case, required education is defined as the mode level of education for the flow of newly hired workers (tenure less than 12 months) in a three-digit occupation in a given year.

Finally, the third measure is defined according to a job analysis based on the Portuguese Classification of Occupations (CPP2010) aggregated at the one-digit level and the International Standard Classification of Education (*ISCED*97). As at the one-digit level the Portuguese Classification of Occupations matches the International Standard Classification of Occupations (*ISCO*08), from now on we will refer to the *ISCO*08 classification.

According to *ISCO*08, occupations at the one-digit level can be divided into ten major groups. For each one-digit level major group, professional job analysts defined a

[5] Given that the distinction between graduate, master, and doctoral degree is reported in *QP* only after 2005 and to use an homogenous criterion over the entire period of 1995-2012, we decided to assign 16 years of schooling for graduates, masters, and doctorates.

[6] As the Portuguese Classification of Occupations (CPP) changed in 2010, the occupation codes valid before 2010 were recoded according to the new classification of occupations. For a few occupations it was not possible to accurately match the codes and, thus, these occupations were excluded from our analysis.

level of competences according to the complexity of tasks and the duties to be performed in each occupation. *ISCO*08 includes four levels of competences that are then related to the level of education required for each competence according to *ISCED*97 (see Figure 1.1 for details).[7] Major groups 1 to 3 are associated with the levels of competences 3 and 4 and require a terciary degree. Major groups 4 to 8 are associated with the level 2 of competences requiring a third cycle of basic education or the secondary. Finally, major group 9 corresponds to the competences level 1 and requires a first or second cycle of basic education. Given that each major group except 3 and 9, is related to more than one level of education, we decided to assign, for each major group, the minimum years of schooling as the required one for each one-digit occupation. The main drawback of this approach is to assume that to all jobs in the same major group correspond the same educational requirements.

*INSERT FIGURE 1.1 HERE*

Based on these definitions of required education and in order to classify the individuals as over- or undereducated, required education for a given occupation is compared to the current level of schooling attained by the worker in that same occupation. Thus, if the attained level of schooling is higher (lower) than the required level of schooling the individual is classified as overeducated (undereducated), otherwise he/she is classified as adequately educated.

Finally, we recognize, as claimed in previous literature, that none of these definitions is exempt from criticism.[8] Having in mind the pros and cons of each measure and the information complementarities among them, throughout this study, and for comparison purposes, we will present the empirical results for the three criteria.

[7] The most recent *ISCED* classification – *ISCED*2011 - does not differ much from *ISCED*97. The main difference arises from a more detailed disagregation of terciary education in the former when compared with the previous version of *ISCED*97. However, as *QP* only discriminates between terciary education levels since 2006, we decided to use the *ISCED*97 classification.

[8] For a detailed discussion of the advantages and limitations of each approach see, for example, Kiker et al. (1997) and Hartog (2000).

### 1.2.3 Incidence of Over- and Undereducation

Table 1.1 reports the average percentage of overeducated and undereducated workers in Portugal for the 1995-2012 period. Using based-data measures, our results show that more than 50 percent of the Portuguese workers in the private sector experience an educational job-mismatch. This percentage increases to 75 percent when we consider an exogenous criterion based on job analysis. This latter result is not unexpected and is in accordance to the value of 70.6 percent reported by Kiker *et al.* (1997) using *QP* data for 1991. Portugal lags behind most european countries with respect to the level of educational attainment of its population (OECD, 2017). In particular, when an exogenous international criterion is used we found a much larger proportion of undereducated workers when compared to a statistical measure (i.e., the mode) that takes into account the distribution of education in Portugal. Finally, as expected, the percentage of overeducated is lower and the percentage of undereducated is higher when we consider the mode of recently hired workers in a three-digit occupation, rather than the mode level of education for the stock of workers.

Regarding the incidence of the phenomenon by gender, the three measures show that, unconditionally, women are more likely to be adequately educated than men. Data-based measures show that males are slightly more likely to be overeducated, while the exogenously-defined measure shows that male workers are more likely to be undereducated and less likely to be overeducated when compared with female workers.

*INSERT TABLE 1.1 HERE*

Figures 1.2 and 1.3 show the same percentages of overeducated and undereducated workers by year and criterion. Overall, on one hand, the data reveal an increasing trend in overeducation and a decreasing trend in undereducation according to the job analysis approach and, on the other hand, a decreasing trend in overeducation and an increasing trend in undereducation according to the statistical measures (in particular from 2003 onwards according to the mode of the stock of employees). These patterns are consistent with a large increase in the levels of educational attainment experienced by the portuguese population in the last two decades and an upgrade of the recruitment educational standards of labor demand. The data suggest that the labor market demand keeped in pace with the rising in the educational attainment of the portuguese population in the past two decades. Notice that, between 1995 and 2012 and based on the mode of newly hired workers, the percentage of undereducated workers almost doubled, i.e., increased from 24.2 percent in 1995 to 40.5 percent in 2012. On the other hand, based on an intemporal job analysis, the percentage of undereducated workers decreased from 65.1 percent in 1995 to 36 percent in 2012.

*INSERT FIGURES 1.2 AND 1.3 HERE*

## 1.3 Econometric Model

### 1.3.1 Two-Way High-Dimensional Fixed Effects Model

To evaluate the impact of educational mismatches on wages we follow the model introduced by Duncan and Hoffman (1981) - the so-called ORU specification. Building up on a standard mincerian wage equation, the authors decompose the attained level of education ($AE$) into three components:

$$AE \equiv RE + OE - UE$$

where $RE$ corresponds to the job required level of education, $OE$ to the number of years of surplus education, and $UE$ to the number of years of deficit education defined as:

$$OE = \begin{cases} AE - RE & \text{if } AE > RE \\ 0 & \text{otherwise} \end{cases}$$

$$UE = \begin{cases} RE - AE & \text{if } RE > AE \\ 0 & \text{otherwise} \end{cases}$$

As in previous research, we acknowledge that observed characteristics of the individual such as education and experience imperfectly reflect their true produtivity (Abowd, Kramarz, and Margolis, 1999; Iranzo, Shivardi, Tosetti, 2008). Unobserved attributes such as ability, motivation, resilience, play an important role in wage determination that, if ignored, lead to biased OLS estimates. Furthermore, we claim that mobility within and across firms in searching for a well-matched job raise endogenous problems, whereas sorting based on unobserved characteristics of both sides of the labor market - workers and firms - is likely to occur (e.g., Baptista *et al.*, 2013; Dahl and Klepper, 2015; Rocha *et al.*, 2018).

Hence, the econometric model that we use is the standard wage equation proposed by Duncan and Hoffman augmented to include controls for worker observed and unobserved (permanent) heterogeneity and firm observed and unobserved (permanent) heterogeneity. The model writes as:

$$\ln w_{ijt} = \alpha_i + \theta_j + \gamma_t + \beta_r RE_{ijt} + \beta_o OE_{ijt} + \beta_u UE_{ijt} + \delta X_{ijt} + \varepsilon_{ijt} \tag{1.1}$$

where $w_{ijt}$ represents the hourly wages (in real euros) for each individual $i$, employed at firm $j$ in year $t$. Hourly wages correspond to total regular payroll (base wage and regular payments) over normal hours worked in the reference month.[9] $\alpha_i$ is a worker fixed effect, $\theta_j$ corresponds to the firm fixed effect, and $\gamma_t$ a time fixed effect. $RE$ is the number of years of required education of the job holding by individual $i$ in year $t$ defined according to the three criteria discussed in Section 1.2, $OE$ and $UE$ are, respectively, surplus and deficit years of education in the job. Vector $X_{ijt}$ includes a quadratic term in the individual's tenure, an age square, and a set of dummies for qualification levels (Table 1.2 presents the descriptive statistics of all the variables introduced in the model, including attained education). $\varepsilon_{ijt}$ is a random error term assumed to be uncorrelated with the regressors.

*INSERT TABLE 1.2 HERE*

### 1.3.2 Wage Returns to Over- and Undereducation: Identification Issues

The parameters of interest are $\beta_r, \beta_o$, and $\beta_u$ corresponding, respectively, to the marginal rate of returns to adequate education, to overeducation, and to undereducation. According to the human capital theory (Becker, 1964), worker productivity is not affected by the job requirements. In this setup, years of required, over-, and undereducation should have a similar return as firms adjust wages to worker's marginal productivity, i.e., $\beta_r = \beta_o = -\beta_u$. On the opposite side, the job competition theory (Thurow, 1975) argues that wages are solely determined by the requirements of the job, restricting the returns to over- and undereducation to be zero, i.e., $\beta_o = \beta_u = 0$. This

[9] Wages were converted into 2010 constant prices using the Consumer Price Index (CPI).

approach assumes that the main competences necessary to perform a job are acquired through on-the-job training. Thus, individuals compete for jobs based on their relative training costs and education serves solely to signal the most able and productive, i.e., the ones that are in a better position in the hiring opportunities queue to get the job as employers need to invest less on-the-job training.

In the midlle ground, the job assignment interpretation (Sattinger, 1993) reconciles both theories arguing that productivity depends on the interaction between job characteristics and worker characteristics, implying that workers with the same educational level may perform differently depending on the job they are doing. According to assignment theory, productivity is maximised when the most skilled are assigned to the most complex and demanding jobs and the least skilled to the simplest jobs. In this context, wages depend on both - job and worker characteristics - restricting the coefficients to $\beta_o > 0$ and $\beta_u < 0$.

In a fixed effects approach, identification of the parameters of interest comes from individuals that changed their educational status in the period under scrutiny. According to the definitons above of $OE$ and $UE$, a switch in the mismatch status takes place if:

i) required education ($RE$) varies as individuals switch occupation within or across firms;

ii) required education ($RE$) varies overtime within an occupation; this is only possible for the data-based measures, as in the exogenous-based criterion required education in a one-digit occupation is time invariant in the analysed period;

iii) the attained level of education ($AE$) varies overtime as individuals invest on formal education while participating in the labor market; in our data aproximately one-fourth of the individuals increased their years of schooling in the analysed period.

Our data show that, regardless the criterion used to identify educational mismatches, at least one-third of the individuals experienced a change in his/her job mismatch status in the 1995-2012 period.[10]

[10] According to the definition based on the mode of all workers, 37.3 percent of the individuals changed their years of surplus education and 30.6 percent changed their years of deficit education at least once in the analysed period. Based on the mode of newly hired workers, 33.1 and 41.6 percent of the individuals changed, respectively, their years of surplus and deficit education. Finally, and according to the job analysis criterion, 27.3 percent changed their years of surplus education and 34.6

## 1.4 Empirical Results

### 1.4.1 Wage Returns to Over- and Undereducation

Table 1.3 (column 3) reports the estimated returns to education based on the full specification defined in equation (1) for the full sample. For comparison purposes, we also provide the estimates of two additional specifications. The first, reported in column (1), replicates the Duncan and Hoffman model and was estimated by OLS. The second specification, reported in column (2), is augmented to account for worker fixed-effects (FE). Finally, the specification in column (3) includes both worker and firm fixed-effects.[11]

The OLS estimates of the returns to education reported in column (1) are consistent with earlier literature. Overeducated workers earn more and undereducated workers earn less than their co-workers adequately educated. Depending on the criterion, the wage gap lies between 4.1 and 4.9 percent for each additional year of surplus education and between 5.1 and 6.2 percent for each extra year of deficit education.[12] The return to required education is around 7.3-7.6 percent. These estimates are very similar to the ones reported by Kiker *et al.* (1997) using the same data set for the year of 1991 and the modal measure of the stock of employed workers: 0.076, 0.048, and $-0.056$ for the coefficients of $RE$, $OE$, and $UE$, respectively.

The worker FE estimates indicate that educational mismatches are correlated with individual unobserved heterogeneity. The results reported in column (2) show a significant reduction in the returns to education once unobserved individual heterogeneity is taken into account, corroborating previous studies using panel data (e.g., Bauer, 2002; Frenette, 2004; Korpi and Tahlin, 2009; Tsai, 2010; Mavromaras *et al.*, 2013). Now

percent their years of deficit education. Notice that in the latter case, changes in the mismatch status solely occur through switches in the one-digit occupation, as required education for a given occupation is fixed overtime, or through the improvements in educational attainment.

[11] The two-way high-dimensional fixed effects models were estimated using the "reghdfe" stata command developed by Sérgio Correia (2016). See also Guimarães and Portugal (2010) for a detailed description of the procedure that allows estimation of a wage equation that includes two high-dimensional fixed effects.

For sake of parsimony, only the estimated coefficients for $RE$, $OE$, and $UE$ are reported in the Tables. Full results are available upon request.

[12] The exact values of the gap are computed as $(exp(\widehat{\beta}) - 1) * 100$.

each extra year of overeducation increases wages by solely 0.7-0.8 percent, while each extra year of undereducation decreases wages by no more than 1.2-1.3 percent.

Finally, once firm permanent observed and unobserved heterogeneity is accounted for, the wage gap between mismatched workers and well-matched workers is further reduced. The estimates in column (3) indicate a return to additional schooling beyond the required level that ranges between 0.4 and 0.5 percent and a penalty to an extra year of deficit schooling relative to the required level that ranges between 0.8 and 0.9 percent, depending on the criterion.[13]

Overall, these results imply that the negative effects of vertical educational mismatches on wages may be largely due to ommited variable bias and self-selection, thereby underlying the need to properly control for unobserved individual heterogeneity. Controlling for heterogeneity in firm paying policies, in addition to worker heterogenity, is also relevant.

*INSERT TABLE 1.3 HERE*

In order to test whether the economics returns to education evolved differently over time, the full-specification reported in column (3) was estimated including an interaction term between the variables $RE$, $OE$, and $UE$ and a dummy variable that takes the value one for the 2004-2012 period. The estimates in Table 1.4 show that the economic returns to education are larger in the most recent period of 2004-2012. This difference is particularly pronounced for the wage returns to an extra year of surplus education, that were almost non-existent in the 1995-2003 period after controlling for worker and firm unobserved heterogeneity, and reach to 0.9-1.1 percent in 2004-2012. The wage penalties to undereducation more than doubled over the 2004-2012 period. For those workers adequately educated the wage returns are also higher in the most recent period.

In sum, these results seem to suggest that the increase in the educational attainments of the portuguese population in the past two decades lead employers to benefit the

[13]The same specifications were estimated separately for the samples of male and female workers (results available upon request). The estimates are quite identical to the ones reported here for the full sample. After accounting for worker and firm fixed effects, the positive return for overeducation lies between 0.3 and 0.4 for males and 0.4 and 0.5 for females. The negative return to undereducation varies between 0.6 and 0.8 percent for males and between 0.9 and 1.0 percent for females. Thus, the wage gap between mismatched and well-matched workers is slightly pronounced for women when compared to men.

overeducated and adequately educated workers and, on the opposite side, to penalize their undereducated counterparts.

*INSERT TABLE 1.4 HERE*

### 1.4.2 The Empirical Distribution of Worker Fixed Effects

Figure 1.4 displays the empirical distribution of permanent worker observed and unobserved heterogeneity for the population of overeducated, undereducated, and adequately educated workers. Worker permanent heterogeneity is proxied by the estimates of the worker fixed effect filtered from firm permanent observed and unobserved heterogeneity obtained by the estimation of the full-model defined in equation (1) using the mode of all workers definition. The graph is based on the 4,209,971 estimates of worker fixed effects. A worker with a higher fixed effect is an individual with a higher earnings premium after controlling for his/her observed attributes, including the eventual educational mismatch, and firm observed and unobserved permanent heterogeneity. It approximates the individual's unobserved time-invariant attributes that are likely to affect productivity and wages.

The graph shows that the empirical distribution of the worker fixed effect for undereducated workers is more shifted to the right, while for the group of overeducated workers is more shifted to the left. The empirical distribution of adequately matched workers lies between the two previous distributions.[14] In other words, undereducated workers seem to correspond to a higher-ability group of employees, while overeducated workers seem to correspond to a low-ability group of employees. The graphs (not reported here) using the fixed effects estimates obtained by estimating the full model using the other two criteria exibhit a similar pattern.

This evidence reinforces the hypothesis that overeducation may emerge as a mechanism to compensate any relative disadvantage in terms of skills, while undereducation may coexist if individuals with low levels of formal education compensate this disadvantage with other forms of human capital relevant to perform the job (e.g., ability).

[14] We also found a negative (positive) coefficient of correlation between the three measures of overeducation (undereducation) and the worker fixed effects that ranges between 6.5 and 9.5 percent (6.7 and 10.0 percent).

*INSERT FIGURE 1.4 HERE*

### 1.4.3 Robustness Checks

In this Section we report the estimates of the returns to over- and undereducation using a specification similar to Verdugo and Verdugo (1989). This model includes the attained level of education ($AE$) instead of the required level of education ($RE$), and uses dummy variables to define the mismatch status. The years of required education are defined according to the criteria discussed in Section 1.2. Thus, a worker is classified as overducated if his/her acquired level of education exceeds the required level of the job ($AE > RE$) and is classified as undereducated if his/her acquired level of education falls below the one required for the current job ($AE < RE$). The results are presented in Table 1.5.

The OLS estimates corroborate previous research, overeducated workers earn less than their similar counterparts with the same years of schooling who are well-matched, while undereducated workers earn more. The penalty for the overeducated varies from 6.6 percent based on the exogenous measure to 11 percent in the case of the data-based measure of the mode of the flow of new hires. Undereducated workers earn a bonus that varies between 3.3 percent and 8.7 percent considering, respectively, the exogenous measure and the data-based measures. Once worker and firm heterogeneity are controlled for, these estimates drop dramatically, and in the case of the data-based measures the gap for the undereducated becomes negative even though very close to zero.

Finally, in Table 1.6 we re-estimate the Verdugo and Verdugo model but redefining the dummies for the educational mismatch. In this specification, an individual is classified as overeducated (undereducated) if his/her years of surplus (deficit) education are simultaneously positive according to the three criteria used to define years of over- and undereducation. The results reveal that after including the worker and the firm fixed effects, a statistically significant penalty of 2.8 percent prevails for the overeducated, as well as a statistically significant small bonus of 0.23 percent for the undereducated.

*INSERT TABLES 1.5 AND 1.6 HERE*

## 1.5 Concluding Remarks

Employing a representative administrative dataset that covers the population of wage earners in the private sector in Portugal, this paper provides a comprehensive analysis on the incidence of the over- and undereducation phenomenons, and their impacts on individual wages. The empirical findings emerging from this exercise are four-fold.

First, the data indicate that half of the employees in the private sector in Portugal experience a vertical educational mismatch. This proportion increases to almost three-fourths when an exogenous measure based on the international standard classification of occupations and education is used.

Second, a time trend analysis apparently revealed contradictory patterns regarding the evolution of the incidence of overeducation and undereducation over time. On one hand, and according to the job analysis definition, the incidence of overeducation has been rising whilst the incidence of undereducation has been decreasing. Given the time-invariant nature of this criteria, this evidence is not surprisingly and merely reflects the huge increase in the attainment schooling of the portuguese labour force in the last two decades. On the other hand, according to the statistical measures based on realized matches, there is indication that overeducation is decreasing and undereducation is increasing as a consequence of the upgrade in employers hiring standards.

Third, accounting simultaneously for worker and firm permanent observed and unobserved heterogeneity reduces dramatically the returns to over- and underducation, suggesting that educational mismatches are largely driven by unobserved characteristics of the worker and the firm and failure to account for them bias the estimates of the mismatch educational effects.

Fourth, the data indicate that the undereducated seem to correspond to a higher-ability group of employees, while the overeducated seem to correspond to a lower-ability group of workers.

# Bibliography


[1] Abowd, J. M., Kramarz, F., Margolis, D. (1999). "High wage workers and high wage firms". Econometrica, 67(2): 251-334.

[2] Alba-Ramirez, A. (1993). "Mismatch in the Spanish labor market: Overeducation?". Journal of Human Resources, 27(2): 259-278.

[3] Bauer, T.K. (2002). "Educational mismatch and wages: a panel analysis". Economics of Education Review, 21: 221-229.

[4] Baptista, R., Lima, F., Preto, M. T. (2013). "Entrepreneurial skills and workers' wages in small firms". Small Business Economics, 40(2): 309-323.

[5] Becker, G. (1964). "Human Capital: A theoretical and empirical analysis with special reference to education". New York: Columbia University Press.

[6] Centeno, M., Novo, A. (2012). "Excess Worker Turnover and Fixed-Term Contracts: Causal Evidence in a Two-Tier System". Labour Economics, 19(3): 320-328.

[7] Chevalier, A. (2003). "Measuring over-education". Economica, 70(279): 509-531.

[8] Chevalier, A., Lindley, J. (2009). "Overeducation and the skills of UK graduates". Journal of the Royal Statistical Society, 172(2): 307-337.

[9] Cohn, E., Khan, S. P. (1995). "The wage effect of overschooling revisited". Labour Economics, 2(1): 67-76.

[10] Cohn, E., Ng, Y. C. (2000). "Incidence and wage effects of overschooling and underschooling in Hong Kong". Economics of Education Review, 19 (2): 159-168.

[11] Correia, S. (2016). "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects". Statistical Software Components s457874, Boston College Department of Economics, revised 25 Jul 2015.

[12] Dahl, M. S., Klepper, S. (2015). "Whom do new firms hire?". Industrial and Corporate Change, 24(4): 819-836.

[13] Dolton, P., Vignoles. A. (2000). "The incidence and effects of overeducation in the U.K. graduate labor market". Economics of Education Review, 19(2): 179-198.

[14] Duncan, G. J, Hoffman, S. D. (1981). "The incidence and wage effects of overeducation". Economics of Education Review, 1(1): 75-86.

[15] Erdogan, B., Bauer, T. N., Peiró, J. M., Truxillo, D. M. (2011). "Overqualified employees: Making the best of a potential bad situation for individuals and organizations". Industrial and Organizational Psycology, 4(2): 215-232.

[16] Figueiredo, H., Teixeira, P., Rubery, J. (2013). "Unequal futures? Mass higher education and graduates' relative earnings in Portugal, 1995-2009". Applied Economics Letters, 20(10): 991-997.

[17] Fine, S. (2007). "Overqualification and selection in leadership training". Journal of Leadership and Organizational Studies, 14(1): 61-68.

[18] Fine, S., Nevo, B. (2008). "Too smart for their own good? A study of perceived cognitive overqualification in workforce". The International Journal of Human Resource Management, 19(2): 346-355.

[19] Fonseca, T., Lima, F., Pereira, S. (2018). "Job polarization, technological change and routinization: Evidence for Portugal". Labour Economics, 51: 317-339.

[20] Frenette, M. (2004). "The overqualified Canadian graduate: The role of the academic program in the incidence, persistence, and economic returns to overqualification". Economics of Education Review, 23(1): 29-45.

[21] Greene, F., McIntosh, S. (2007). "Is there a genuine under-utilization of skills amongst the over-qualified?". Applied Economics, 39(4): 427-439.

[22] Groot, W. (1993). "Overeducation and the returns to enterprise-related schooling". Economics of Education Review, 12(4): 299-309.

[23] Groot, W., van den Brink M. H. (2000). "Overeducation in the labor market: A meta-analysis". Economics of Education Review, 19(2): 149-158.

[24] Guimarães. P., Portugal, P. (2010). "A Simple Feasible Alternative Procedure to Fit Models with High-Dimensional Fixed Effects". The Stata Journal, 10(4): 628-49.

[25] Hartog, J. (2000). "Over-education and earnings: Where are we, where should we go?". Economics of Education Review, 19(2): 131-147.

[26] Hartog, J., Oosterbeek, H. (1988). "Education, allocation and earnings in the Netherlands: Overschooling?". Economics of Education Review, 7(2): 185-194.

[27] Hartog, J., Groeneveld, S. (2004). "Overeducation, wages and promotions within the firm". Labour Economics, 11(6): 701-714.

[28] Hersch, J. (1991). "Education match and job match". Review of Economics and Statistics, 73(1): 140-144.

[29] ILO (2014). "Skills mismatch in Europe: statistics brief". International Labour Office (Geneva).

[30] Iranzo, S., Shivardi, F., Tosetti, E. (2008). "Skill dispersion and firm productivity: An analysis with employer-employee data". Journal of Labor Economics, 26(2): 247-285.

[31] Kiker, B. F., Santos, M. C., De Oliveira, M. M. (1997). "Overeducation and undereducation: evidence for Portugal". Economics of Education Review, 16(2): 111-125.

[32] Korpi, T.; Tåhlin, M. (2009). "Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000". Labour Economics, 16(2): 183-193.

[33] Martins, P. (2009). "Dismissals for cause: The difference that just eight paragraphs can make". Journal of Labor Economics, 27(2): 257-279.

[34] Maynard, D. C., Joseph, T. A., Maynard, A. M. (2006). "Underemployment, job attitudes, and turnover intensions". Journal of Organizational Behavior, 27(4): 509-536.

[35] Mavromaras, K., McGuinness, S., O'Leary, N., Sloane, P, Wei, Z. (2013). "Job mismatches and labour market outcomes: Panel evidence on Australian university graduates". Economic Record, 89 (286): 382-395.

[36] McGuinness, S. (2006). "Overeducation in the labor market?". Journal of Economic Surveys, 20(3): 387-418.

[37] Murillo, I. P., Rahona-López, M., Salinas-Jiménez, M. M. (2012). "Effects of educational mismatch on private returns to education: An analysis of the Spanish case (1995-2006)". Journal of Policy Modeling, 34(5): 646-659.

[38] OECD (2012). OECD Economic Surveys: Portugal 2012. OECD Publishing, Paris.

[39] OECD (2017). Education at a glance 2017: OECD Indicators. OECD Publishing, Paris.

[40] Oliveira, M. M., Santos, M. C., Kiker, B. F. (2000), "The role of human capital and technological change in overeducation". Economics of Education Review, 19(2): 199-206.

[41] Quintini, G. (2011). "Over-qualified or under-skilled: A review of existing literature". OECD Social, Employment and Migration WP No 121. DELSA/ELSA/WD/SEM(2011)6.

[42] Rocha, V., Carneiro, A., Varum, C. (2018). "Leaving employment to entrepreneurship: The value of co-worker mobility in pushed and pulled driven start-ups". Journal of Management Studies, 55(1): 60-85.

[43] Rocha, V., van Praag, M., Folta, T., Carneiro, A. (2018, in press). "Endogeneity in strategy-performance analysis: An application to initial human capital strategy and new venture performance". Organizational Research Methods, 1-25.

[44] Rumberger, R.W. (1987). "The impact of surplus schooling on productivity and earnings". Journal of Human Resources, 22(1): 24-50.

[45] Sattinger, M. (1993). "Assignment models of the distribution of earnings". Journal of Economic Literature, 31(2): 831-880.

[46] Sicherman, N. (1991). "Overeducation in the labor market". Journal of Labor Economics, 9(2): 101-122.

[47] Thurow, L.C. (1975). "A job-competition model", In: Priore, M. J. (Ed.): *Unemployment and inflation: Institutionalist and Structuralist Views.* New York, 17-32.

[48] Tsai, Y. (2010). "Returns to overeducation: A longitudinal analysis of the U.S. labor market". Economics of Education Review, 29(4): 606-617.

[49] Tsang, M. C., Levin, H. M. (1985). "The Economics of Overeducation ". Economics of Education Review, 4(2): 93-104.

[50] Tsang, M. C., Rumberger, R. W., Levin, H. M. (1991). "The impact of surplus schooling on worker productivity". Industrial Relations, 30(2): 209-228.

[51] Van Smoorenburg, M., Van der Velden, R. (2000). "The training of school leavers, complementarity or substitution?". Economics of Education Review, 19(2): 207-217.

[52] Verdugo, R., Verdugo, N. (1989). "The impact of surplus schooling on earnings". Journal of Human Resources, 24(4): 629-643.

# Tables and Figures

Figure 1.1 – Required education: Job Analysis based on ISCO08 and ISCED97

| 1-Digit Level | Major group according to ISCO08 | Competences level | ISCED97 Groups | Years of schooling |
|---|---|---|---|---|
| **1** | Representative of the state legislature and executive bodies, executives, directors, executive managers. | 3 | 5b Bachelor<br>5a Graduate and Post-graduate | **15** |
| | | 4 | 6 Master and Doctorate<br>Second stage of tertiary education | |
| **2** | Expert on intellectual and scientific activities | 4 | 5a Graduate and Post-graduate<br>6 Master and Doctorate | **16** |
| **3** | Technicians and associate professionals | 3 | 5b Bachelor | **15** |
| **4** | Administrative staff | 2 | 2 Third Cycle of basic education<br>3 and 4 Secondary | **9** |
| **5** | Service and sales workers | 2 | 2 Third Cycle of basic education<br>3 and 4 Secondary | **9** |
| **6** | Skilled agricultural, forestry and fishery workers | 2 | 2 Third Cycle of basic education<br>3 and 4 Secondary | **9** |
| **7** | Skilled construction and industry sector workers | 2 | 2 Third Cycle of basic education<br>3 and 4 Secondary | **9** |
| **8** | Plant and machine operators and assemblers | 2 | 2 Third Cycle of basic education<br>3 and 4 Secondary | **9** |
| **9** | Unskilled workers | 1 | 1 First and second Cycle of basic education | **4** |

Citing ISCO08: Occupations at Level 1 "involve the performance of simple and routine physical or manual tasks"; at Level 2 "involve the performance of tasks such as operating machinery and electronic equipment, driving vehicles; maintenance and repair of electrical and mechanical equipment; and manipulation, ordering and storage of information; at Level 3 "involve the performance of complex technical and practical tasks that require an extensive body of factual, technical and procedural knowledge in a specialized field; at Level 4 "involve the performance of tasks that require complex problem-solving, decision-making and creativity based on an extensive body of theoretical and factual knowledge in a specialized field.

**Sources: ISCO08 and ISCED97**

Figure 1.2 – Incidence of overeducation by criteria, Portugal 1995-2012

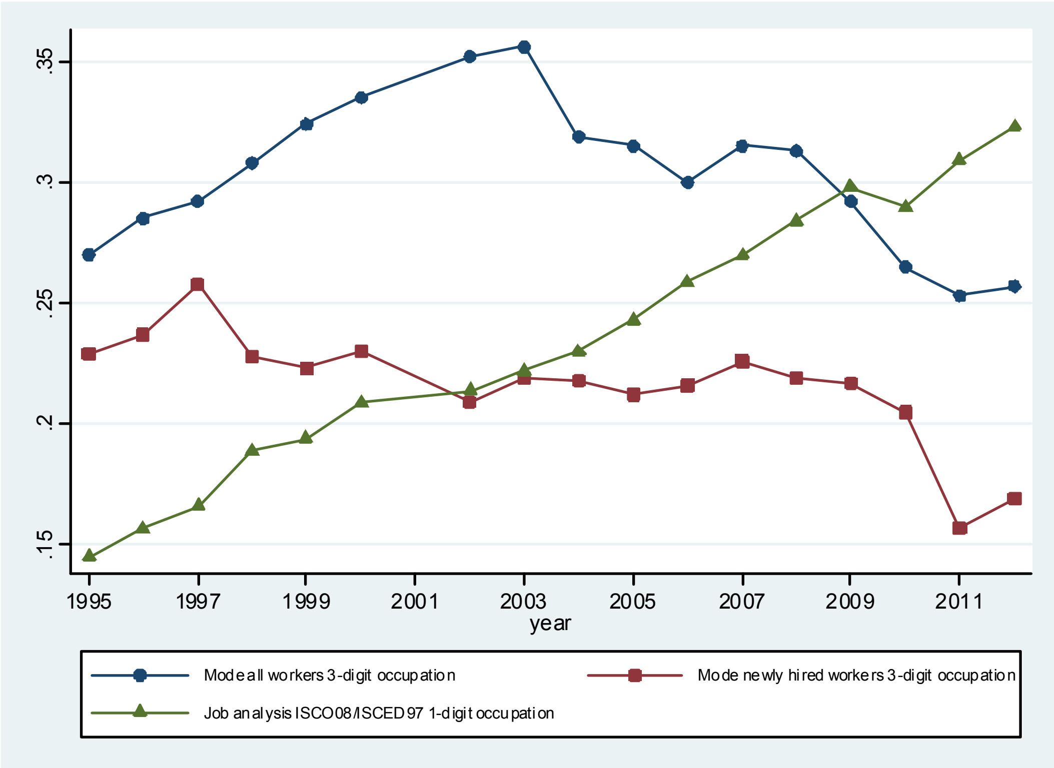

Figure 1.3 – Incidence of undereducation by criteria, Portugal 1995-2012

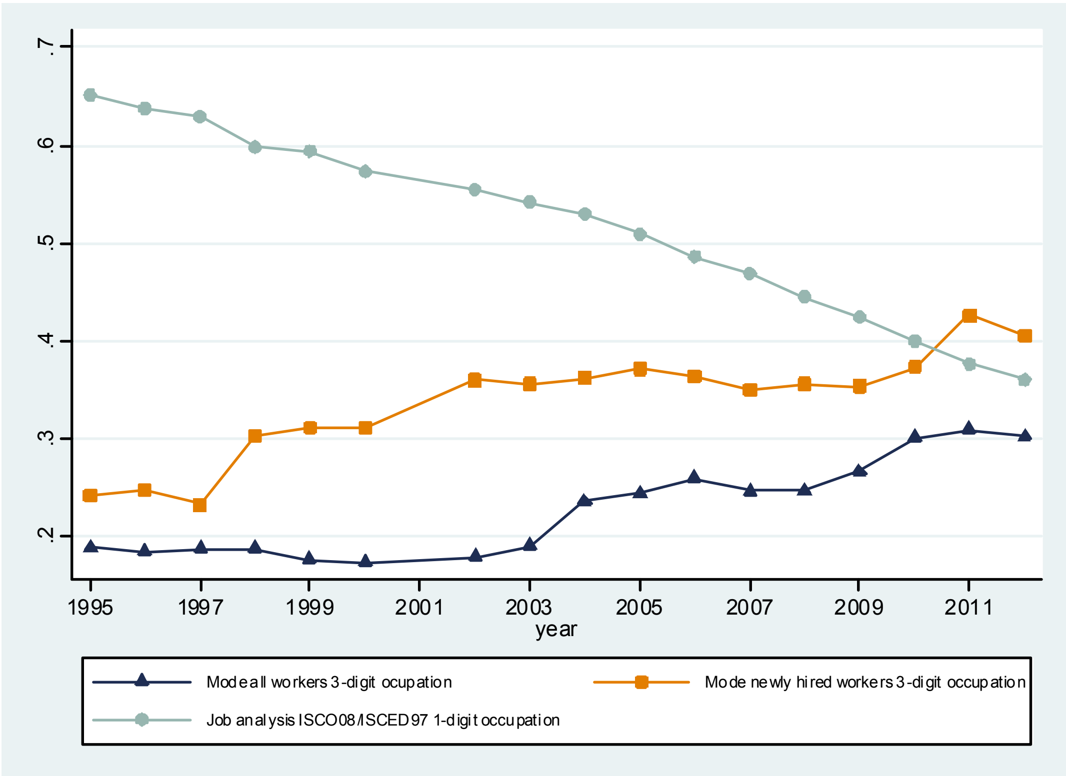

Figure 1.4 – Empirical distribution of worker fixed effects

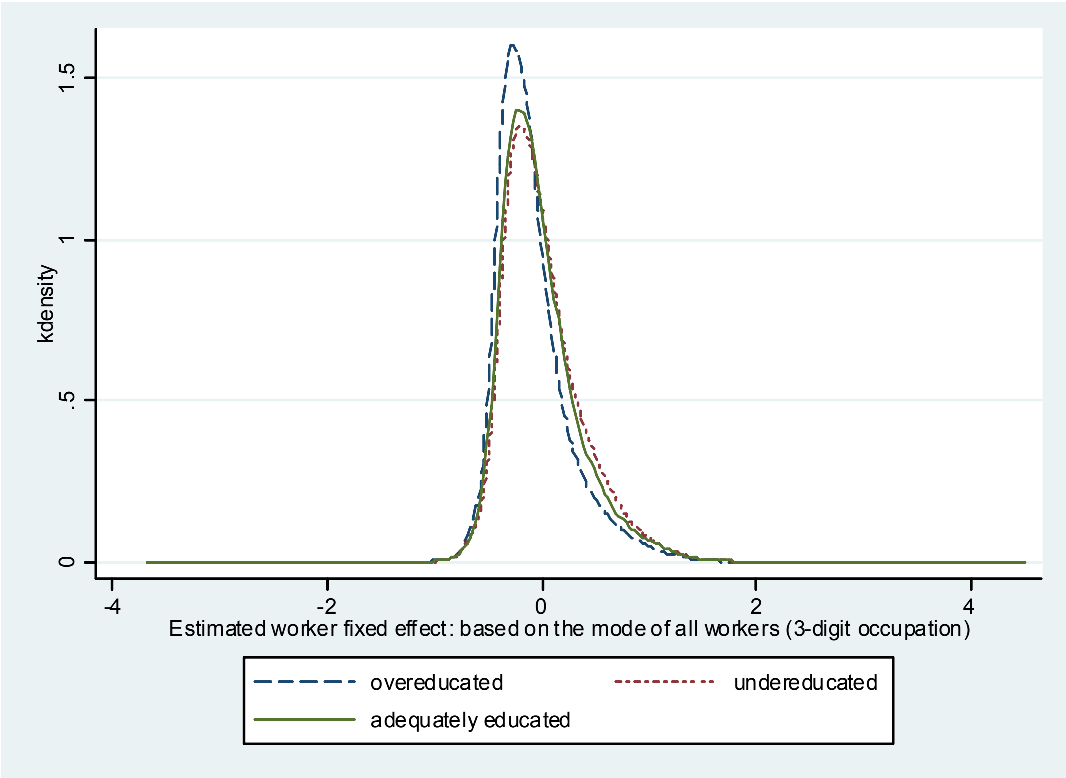

Table 1.1 – Incidence of overeducation and undereducation by criterion, Portugal 1995-2012

| | Full sample | Males | Females |
|---|---|---|---|
| | N=27,268,443 | N=15,981,429 | N=11,287,014 |
| Mode all workers (3-digit occupation) | | | |
| Overeducated (%) | 30.1 | 32.1 | 27.4 |
| Undereducated (%) | 23.7 | 23.4 | 24.1 |
| Mode newly hired workers (3-digit occupation) | | | |
| Overeducated (%) | 21.2 | 22.3 | 19.7 |
| Undereducated (%) | 34.7 | 35.5 | 33.6 |
| Job analysis ISCO08/ISCED97 (1-digit occupation) | | | |
| Overeducated (%) | 24.7 | 19.7 | 31.7 |
| Undereducated (%) | 49.8 | 57.3 | 39.1 |

Table 1.2 – Descriptive statistics, Portugal 1995-2012

$N = 27,268,443$

| | Mean | St. Dev. | Min. | Max. |
|---|---|---|---|---|
| Log real hourly earnings (in 2010 euros) | 1.55 | 0.53 | -0.56 | 9.72 |
| Female | 0.41 | | 0 | 1 |
| Age (years) | 38.00 | 10.79 | 15 | 64 |
| Tenure (months) | 98.07 | 102.89 | 0 | 656 |
| Qualification levels | | | | |
| Top executives | 0.08 | | 0 | 1 |
| Intermediary executives | 0.05 | | 0 | 1 |
| Supervisor, team leader, foreman | 0.04 | | 0 | 1 |
| High-skilled professionals | 0.07 | | 0 | 1 |
| Skilled professionals | 0.42 | | 0 | 1 |
| Semi-skilled professionals | 0.18 | | 0 | 1 |
| Non-skilled professionals | 0.12 | | 0 | 1 |
| Apprentices, interns, trainees | 0.04 | | 0 | 1 |
| Education (years) | 8.21 | 4.07 | 0 | 16 |
| Mode all workers (3-digit occupation) | | | | |
| Required education (years) | 8.05 | 4.12 | 4 | 16 |
| Overeducation (years) | 1.10 | 1.98 | 0 | 12 |
| Undereducation (years) | 0.95 | 1.95 | 0 | 16 |
| Mode newly hired workers (3-digit occupation) | | | | |
| Required education (years) | 8.81 | 3.88 | 4 | 16 |
| Overeducation (years) | 0.75 | 1.64 | 0 | 12 |
| Undereducation (years) | 1.36 | 2.19 | 0 | 16 |
| Job analysis ISCO08/ISCED97 (1-digit occupation) | | | | |
| Required education (years) | 9.59 | 3.43 | 4 | 16 |
| Overeducation (years) | 0.85 | 1.77 | 0 | 12 |
| Undereducation (years) | 2.23 | 2.60 | 0 | 16 |

Table 1.3 – Wage regressions: Duncan and Hoffman model, Portugal 1995-2012

| | Dependent variable: log real hourly wages | | |
|---|---|---|---|
| | OLS | Worker FE | Worker & Firm FE |
| | (1) | (2) | (3) |
| Mode all workers | | | |
| $RE$ | 0.0733*** | 0.0149*** | 0.0108*** |
| | (0.00006) | (0.00009) | (0.00008) |
| $OE$ | 0.0421*** | 0.0077*** | 0.0049*** |
| | (0.00008) | (0.00008) | (0.00007) |
| $UE$ | −0.0508*** | −0.0128*** | −0.0089*** |
| | (0.00010) | (0.00009) | (0.00008) |
| $R^2$ overall / $R^2$ within | 0.546 | 0.882/0.1748 | 0.909/0.1617 |
| Mode newly hired workers | | | |
| $RE$ | 0.0757*** | 0.0142*** | 0.0098*** |
| | (0.00006) | (0.00009) | (0.00008) |
| $OE$ | 0.0408*** | 0.0066*** | 0.0039*** |
| | (0.00010) | (0.00009) | (0.00008) |
| $UE$ | −0.0509*** | −0.0128*** | −0.0090*** |
| | (0.00009) | (0.00009) | (0.00008) |
| $R^2$ overall / $R^2$ within | 0.546 | 0.882/0.1745 | 0.909/0.1613 |
| Job analysis ISCO08/ISCED97 | | | |
| $RE$ | 0.0762*** | 0.0177*** | 0.0140*** |
| | (0.00008) | (0.00009) | (0.00008) |
| $OE$ | 0.0489*** | 0.0071*** | 0.0041*** |
| | (0.00010) | (0.00010) | (0.00008) |
| $UE$ | −0.0624*** | −0.0116*** | −0.0081*** |
| | (0.00008) | (0.00009) | (0.00008) |
| $R^2$ overall / $R^2$ within | 0.536 | 0.882/0.1765 | 0.909/0.1638 |
| $N$ | 27,268,443 | 27,042,824 | 26,646,372 |

Notes: (i) controls include age, age$^2$, tenure, tenure$^2$, gender dummy, qualification dummies, and time dummies; In the worker FE and the worker & firm FE model, controls include age$^2$, tenure, tenure$^2$qualification and time dummies;
(ii) worker-cluster robust standard errors in parentheses;
(iii) *** denote significant at 1 percent.

Table 1.4 – Wage regressions: Duncan and Hoffman model with time period interactions, Portugal 1995-2012

| Dependent variable: log real hourly wages<br>Full-specification: Worker & Firm FE | Mode all workers<br>(1) | Mode newly hired workers<br>(2) | Job analysis ISCO08/ISCED97<br>(3) |
|---|---|---|---|
| $RE$ | 0.0045*** | 0.0038*** | 0.0068*** |
| | (0.00008) | (0.00008) | (0.00009) |
| $OE$ | 0.0006*** | -0.0002* | -0.0023*** |
| | (0.00008) | (0.00009) | (0.00011) |
| $UE$ | $-0.0040$*** | $-0.0046$*** | $-0.0034$*** |
| | (0.00010) | (0.00009) | (0.00009) |
| $RE \times Y2004\_2012$ | 0.0111*** | 0.0120*** | 0.0114*** |
| | (0.00005) | (0.00005) | (0.00006) |
| $OE \times Y2004\_2012$ | 0.0085*** | 0.0088*** | 0.0109*** |
| | (0.00008) | (0.00009) | (0.00010) |
| $UE \times Y2004\_2012$ | $-0.0061$*** | $-0.0054$*** | $-0.0074$*** |
| | (0.00009) | (0.00008) | (0.00007) |
| $R^2$ overall / $R^2$ within | 0.910/0.1684 | 0.910/0.1683 | 0.910/0.1698 |
| $N$ | 26,646,372 | 26,646,372 | 26,646,372 |

Notes: (i) controls include age$^2$, tenure, tenure$^2$, qualification dummies;
(ii) worker-cluster robust standard errors in parentheses;
(iii) *, **, *** denote significant at 10, 5 and 1 percent, respectively

Table 1.5 – Wage regressions: Verdugo and Verdugo model, Portugal 1995-2012

| Dependent variable: log real hourly wages | | | |
|---|---|---|---|
| | OLS | Worker FE | Worker & Firm FE |
| | (1) | (2) | (3) |
| Mode all workers | | | |
| $AE$ | 0.0687*** | 0.0123*** | 0.0084*** |
| | (0.00006) | (0.00009) | (0.00007) |
| $OE(=1\ if\ RE < AE)$ | −0.0990*** | −0.0184*** | −0.0153*** |
| | (0.00035) | (0.00025) | (0.00022) |
| $UE(=1\ if\ RE > AE)$ | 0.0875*** | −0.00007 | −0.0011*** |
| | (0.00039) | (0.00025) | (0.00022) |
| $R^2$ overall / $R^2$ within | 0.541 | 0.882/0.1738 | 0.909/0.1609 |
| Mode newly hired workers | | | |
| $AE$ | 0.0712*** | 0.0121*** | 0.0080*** |
| | (0.00006) | (0.00009) | (0.00008) |
| $OE(=1\ if\ RE < AE)$ | −0.1085*** | −0.0189*** | −0.0141*** |
| | (0.00038) | (0.00026) | (0.00023) |
| $UE(=1\ if\ RE > AE)$ | 0.0874*** | −0.00009 | −0.0029*** |
| | (0.00035) | (0.00022) | (0.00020) |
| $R^2$ overall / $R^2$ within | 0.541 | 0.882/0.1737 | 0.909/0.1608 |
| Job analysis ISCO08/ISCED97 | | | |
| $AE$ | 0.0692*** | 0.0145*** | 0.0106*** |
| | (0.00007) | (0.00009) | (0.00008) |
| $OE(=1\ if\ RE < AE)$ | −0.0663*** | −0.0241*** | −0.0203*** |
| | (0.00048) | (0.00035) | (0.00030) |
| $UE(=1\ if\ RE > AE)$ | 0.0325*** | 0.0204*** | 0.0194*** |
| | (0.00044) | (0.00032) | (0.00028) |
| $R^2$ overall / $R^2$ within | 0.531 | 0.882/0.1747 | 0.909/0.1619 |
| $N$ | 27,268,443 | 27,042,824 | 26,646,372 |

Notes: (i) controls include age, age$^2$, tenure, tenure$^2$, gender dummy, qualification dummies, and time dummies;
(ii) worker-cluster robust standard errors in parentheses;
(iii) ***, **, * denote significant at 1, 5, and 10 percent, respectively.

Table 1.6 – Wage regressions: Verdugo and Verdugo model, Portugal 1995-2012 - alternative definition of over- and undereducation

| | Dependent variable: log real hourly wages | | |
|---|---|---|---|
| | OLS | Worker FE | Worker & Firm FE |
| | (1) | (2) | (3) |
| Mode all workers | | | |
| $AE$ | 0.0714*** | 0.0128*** | 0.0089*** |
| | (0.00006) | (0.00008) | (0.00007) |
| $OE$ | −0.1302*** | −0.0327*** | −0.0281*** |
| | (0.00050) | (0.00035) | (0.00031) |
| $UE$ | 0.1092*** | 0.0048*** | 0.0023*** |
| | (0.00042) | (0.00026) | (0.00023) |
| $R^2$ overall / $R^2$ within | 0.539 | 0.882/0.1741 | 0.909/0.1613 |
| $N$ | 27,268,443 | 27,042,824 | 26,646,372 |

Notes: (i) controls include age, age$^2$, tenure, tenure$^2$, gender dummy, qualification dummies, and time dummies; In the worker FE and the worker & firm FE models, controls include age$^2$, tenure, tenure$^2$, qualification and time dummies;
(ii) worker-cluster robust standard errors in parentheses;
(iii) *** denote significant at 1 percent;
(iv) $OE = 1$ if overeducated by the three criteria simultaneously;
(iv) $UE = 1$ if undereducated by the three criteria simultaneously.

# Chapter 2

# Overqualification and Future Labor Market Outcomes: Evidence from Recent Graduates in Portugal

Abstract: Using a sample of higher education graduates in Portugal, this paper analyses the career dynamics of workers who entered the labor market for the first time in a job for which they were overqualified. Exploring a large matched employer-employee data set over the 2006-2012 period, we provide evidence that overqualification is a persistent phenomenon where workers get trap. Six years after entering the labor market, 63% of the workers that entered overqualified remain in that status. Finally, even though, unconditionally, wages at entry are lower for overqualified workers when compared with the wages of well-matched workers, overqualified workers that were able to switch to a well-matched job over the analysed period exhibit a larger wage growth. Actually, taking into account workers observed and unobserved permanent heterogeneity, the estimates reveal that five years after entering the labor market overqualified individuals that were able to switch to a well-matched job experience a wage growth that exceeds the wage growth of their similar well-matched counterparts in 12 percentage points.

KEYWORDS: skill mismatches, overqualification, wages, persistence, recent graduates
JEL CODES: I26; J24; J31; J62

## 2.1 Introduction

In the past decades, there has been a growing interest on skill mismatch, more specifically on vertical mismatch.[1] The vast majority of research on skill mismatch has focused on overeducation by confronting attained education levels of workers with the education levels required in the job (e.g., McGuinness *et.al.*, 2017). A branch of this literature analysed the impact of overeducation on wages showing that overeducated workers suffer a penalty when compared to their similar counterparts with the same years of schooling who are well-matched, providing a static view of the phenomenon.

More recent studies, using longitudinal data, provide a dynamic approach of the phenomenon in terms of career patterns of the overeducated/overqualified. This paper is more closely related with this literature that aims to analyse the medium and long-term effects of overeducation/overqualification on future labor market outcomes (e.g., Meroni *et al.*, 2017).

The paper by Tsang and Levin (1985) was one of the first attempts to study the persistence of skill mismatches. Since then, there has been a substantial debate and opposite conclusions on the duration of overeducation and overqualification and its scarring effects on future labor market outcomes.

According to the human capital theory, whose seminal work is from Becker (1964), the labor market is a perfect market where individuals are encouraged to increase their educational levels as in return they will receive a considerable payoff (Freeman, 1976). An over investment in education should not occur, and if it does this mismatch would only be possible on the short-run. Hence, overeducation is seen as a compensation for a lack of other human capital endowments, such as ability or experience, typically affecting recent entrants in the labor market such as young people (Becker, 1964; Groot *et al.* 2000). According to Becker (1964), if a firm employs overeducated workers they will use their skill surplus to train their co-workers, otherwise, they will be remunerated below their potential marginal product as wages are equal to the individual's marginal product (Mincer, 1974).

An alternative theory suggesting that overeducation is a transitory phenomenon

[1] A vertical mismatch occurs when the level of education/qualification is higher or lower than the one required by the job.

is the career mobility theory proposed by Sicherman and Galor (1990). This model predicts that workers may temporarily work in jobs whereby they are overeducated, but that provide them with skills to be used later in a different higher-level job, because the probability of promotion would be higher. Accordingly, workers deliberately choose jobs below their own level of education in order to acquire relevant on-the-job training and experience promoting upward mobility and a fast career progression (Erdsiek, 2017). The mismatch is seen as a kind of investment in work experience that would permit promotions or career opportunities in the future, a stepping stone to a well-matched job (Baert *et al.*, 2013).

In other studies, emphasis has been placed on the higher persistence of overqualification. In this line of thought, two theories stand out: the job competition theory from Thurow (1975) and the job assigment model from Sattinger (1993).

According to the job competition theory, the labor market is imperfect and mismatch is the result of a persistent market failure. This approach emphasizes the role of the demand side of the labor market and stresses the importance of on-the-job training. Labor markets are seen as training markets where firms recruit workers and offer them a training to acquire the skills they need to perform their jobs. In that sense, wages do not depend on the abilities or characteristics of workers, but instead, on the job characteristics. Workers are ranked in an imaginary labor queue where higher educated individuals are in the front of the queue increasing their chances to obtain a better job. Consequently, workers are encouraged to increase their educational level thus enabling them to move up their queue positioning. In this competition between workers to reach the best position in the queue, some graduates accumulate more education than the one required by the job and become overeducated. Furthermore, workers in low level occupations will acquire less skills through on-the-job training and will be more likely to get trap into a mismatch condition (Kiersztyn, 2013). Consequently, overeducation is likely to be a persistent phenomenon.

The assignment theory, whose seminal work is Sattinger (1993), has been considered a middle ground between the human capital theory and the job competition theory. This approach considers that the actual productivity is the maximum of the worker's productivity. A worker is "assigned" to a sector where he will choose the job that maximizes his utility according to his characteristics as well as the earnings. The choice of the sector is an intermediate step between an individual's characteristics and their earnings (McGuinness, 2006). In this framework, educational mismatches may be the consequence of errors in the complicated assignment mechanism (Büchel, 2002) or of imbalances between the supply and demand for individuals with different skills (McGuinness, 2006; Quintini, 2011). These mismatches may be persistent if the labor market does not respond to changes in the supply of workers with different educational levels (Sloane, 2003; Kiersztyn, 2013).

The empirical contribution of this paper to the discussion on whether overeducation/overqualification is a temporary or transitory phenomenon is twofold. First, our work aims to contribute to the literature on the persistence of skill mismatch using a measure of overqualification that takes into account the relative importance of tasks within each occupation at a 2 digit-level based on the O*NET classification. We propose to study whether new entrants graduates remain overqualified throughout their tenure or overqualification is a transitory phenomenon and a stepping-stone to a better job. Literature on the persistence of overqualifiaction is ambiguous, however the consequences of skill mismatches, both at the individual and the economy level, will depend crucially on its longevity (Erdsiek, 2017).

Second, this paper revisits the literature on skill mismatch seeking to examine whether being overqualified in the first job leaves a scar on future wages of recent graduates.[2]

We based our empirical analysis on panel data. We combine two data sets, a rich matched employer-employee dataset (with information on workers, firms, and establishments) and a database containing measures of occupational characteristics and worker requirements information, for the period between 2006 and 2012.

To address the first research question we identifiy and follow recent overqualified graduates at risk of making a transition to a suitable job. Using continuous time and a proportional hazard model that takes into account workers' unobserved heterogeneity, we analyse whether the skill mismatch of recent graduates exhibits a persistent or transitory pattern. Consistent with previous literature, we find that overqualification is a permanent phenomenon where newly graduates workers get trap. At the end of the time span, 53.6% of newly graduates who entered the labor market overqualifed, failed to move to an adequate job according to their skills. Our results show that less able

[2] According to Mavromaras *et al.* (2012), a scarring effect is "a disadvantage that is self-perpetuating for the individual and is clearly over and above any positive or negative effect that the individual characteristics may play regarding the presence or absence of this disadvantage".

workers tend to slow down their transition and that the speed of transition is largely influenced by individuals unobserved heterogeneity.

To analyse the second research question, we estimate a wage equation that allows us to compare the evolution of wages of overqualified workers against similar well-matched workers, by controlling for both workers and firm observed and unobserved permanent heterogeneity. Our results show that entry wages of overqualified workers are lower than entry wages of well-matched workers with similar individual observed characteristics. However, after taking into account worker and firm unobserved permanent heterogeneity, the results indicate that this gap tends to decrease in the first years in the labor market for overqualified workers that were able to move to a well-matched job. For those that remain overqualified in the analysed period the gap tends to prevail.

The paper is organized as follows. Section 2.2 briefly reviews the main empirical literature on the effects and impacts of overqualification on future labor market outcomes. Section 2.3 describes the data and offers a statistical comparison between recent overqualified graduates workers and well-matched workers. Empirical results on the persistence of overqualification and on wages prospects are presented and discussed in Sections 2.4 and 2.5, respectively. Section 2.6 concludes.

## 2.2 Literature Review

Fundamentally, empirical research on overqualification provides two opposite conclusions on the persistence of individual job mismatches. According to early studies, overqualification is a transitory mismatch supporting the hypothesis of the career mobility theory (Sicherman, 1991). The intuition is that overqualification serves as a "stepping-stone" to better future jobs and affects more particularly new entrants who lack work-experience, i.e., workers accept temporarily jobs for which they are overqualified, so they can acquire on-the-job training and experience to progress in their career (see, e.g., Robst, 1995). Sicherman (1991), using data from the Panel Study of Income Dynamics in 1976 and 1978, find that overqualified workers experience a higher turnover and upward mobility. Alba-Ramirez (1993), find evidence in Spain that supports the hypothesis that on-the-job training and experience can provide overeducated workers with the qualifications that match their job market expectations based on years of education. Dorn and Sousa-Poza (2005) also support that spells of overqualification are relatively short in Switzerland. They find that about half of all individuals who were overqualified in a given year had an adequate job match one year later. Frei and

Sousa-Poza (2012), using panel data from the first eight waves of the Swiss Household Panel (SHP), covering the period 1999 to 2006, find that spells of overqualification are relatively short and about half of all individuals who were overqualified in a given year found a well-matched job in the next year.

More recent studies have questioned the validity of the career mobility theory and showed that the duration of overqualification is long lasting. In this line of thought, overqualification is the consequence of labor market frictions (Kiersztyn, 2013). Dolton and Vignoles (2000) find that a significant proportion of U.K. graduates that have left higher education in 1980 entered the labor market overqualified. Six years later, 30% were still in jobs for which they were overqualified. Verhaest and van der Velden (2013), using data for a set of European countries and Japan, suggest that graduates that enter the labor market during a recession are much more likely to be overeducated in their first job affecting the quality of the match up to five years later. Rubb (2003), using the Current Population Surveys (CPS) for U.S., finds that approximately three out of four overqualified individuals in a given year were overqualified in the next year. Büchel and Mertens (2004) find, for Germany, that overqualified workers are less likely to experience an upward occupational mobility. Concerning the persistence of this mismatch, they conclude that the career mobility theory fails to explain the widespread persistence of overqualification observed in Germany. This argument extends to other countries, for example, in Canada, Frenette (2004) finds that more than half of all graduates of master's programs are overqualified in their job, even five years after graduating. For Belgium, Baert *et al.* (2013) find that overeducation is a trap, especially early in the unemployment spell.

Cohort studies of school leavers have focused on workers at the beginning of their career and attempted to test the validity of the career mobility theory. Kiersztyn (2013), for Poland, emphasizes the higher risk of young cohorts of workers to be overqualified when compared to other cohorts, while Verhaest and Schatteman (2010) argue that 30 to 40% of Flemish school leavers who have entered the labor market overeducated remain mismatched most of the time during the seven years that follow their entry. Both studies support the idea that overqualification is a trap for workers, however they fail to explain the reasons of this persistence.

In order to evaluate the effects of true dependence on overqualification, some studies take into account individuals' unobserved heterogeneity.[3] Blàzquez and Budría (2012) found that 17,6% of the observed persistence in overeducation in Germany is a consequence of having been in the same mismatch in the previous year. Boll *et al.* (2016) and Mavromas and McGuinness (2012) corroborate the previous conclusion and found that overqualification exhibits true state dependence, respectively in Germany and Australia.

Recently, emphasis has been placed on the dynamics of overqualification along workers' careers. Meroni *et al.* (2017) provide some empirical evidence regarding the effects of being mismatch at the beginning of a career and the consequence on future career prospects for recent graduates. They find that it is better to wait for a well-matched job rather than accept a job for which recent graduates workers are overeducated at the beginning of their career. Furthermore, overeducation is a trap and overeducated workers are less likely to transit to a position where they are well-matched. For Scandinavian graduates, larger negative effects are found if they are apparently overeducated in their first job. These findings have been confirmed by Clark *et al.* (2017), who pointed out that overeducation is a persistent phenomenon and overeducated workers are less likely to transit to a well-matched job (in particular blacks and low-ability workers), associated with persistent and sizeable lower wages. Once controlling for both observed and unobserved individual heterogeneity they found that true dependence is not so relevant. Rather, the propensity to exit overeducated employment appears to be very heterogeneous among workers. They find that past unemployment is associated with a higher duration of future overeducation spells, thus indicating that overeducation is likely to be one of the mechanisms through which the scarring effects on earnings associated with unemployment spells operate. The scarring effects associated with overeducation could also account for some of the negative wage effects of graduating during a recession which have been recently uncovered in the literature.

Erdsiek (2017), using data for Germany, showed that overqualification is highly persistent among tertiary graduates over the first ten years of their career cycle. The author also finds that unobserved heterogeneity explains the high persistence of this mismatch.

[3] According to Erdsiek (2017), a true state dependence refers to a "genuine behavioural effect of previous overqualification".

## 2.3 Data and Methodological Issues

### 2.3.1 Quadros de Pessoal (QP) and O*NET

We use two datasets in our analysis: *Quadros de Pessoal* (*QP*) and the *Occupational Information Network database* (O*NET, version 21.0).

*QP,* a matched employer-employee dataset collected by the Portuguese Ministry of Labor, Solidarity, and Social Security, is an annual mandatory survey that all Portuguese firms, in the private sector, with at least one wage earner are legally obliged to fill in.[4] The data include information at the firm (e.g., industry, location, employment, sales, ownership), establishment (e.g., location, employment, and economic activity), and worker level (e.g., gender, age, education, occupation, qualification, tenure, wages, hours worked).

All firms, establishments, and workers entering *QP* dataset have a unique identifying number that makes it possible to follow them across all annual sets of data. Furthermore, the worker files include the firm and establishment number to which each individual is affiliated in a given year, allowing to match workers with their employers both at the firm and the establishment level. We restrict our analysis to the period 2006-2012 as the education codes changed in 2006.

The present analysis targets workers having at least a bachelor degree and who entered the labor market for the first time in 2006 or 2007. We therefore select workers that have never appeared in the *QP* files in the previous five years before entry and whose year of admission in the firm at entry is 2006 or 2007, according to the respective cohort. Furthermore, and to assure that we are including in our analysis recent labor market entrants, we excluded from our analysis workers aged above 29 years old at the time of their entry in the labor market.[5] Moreover, and whenever possible, we use some correction routines to repair missing and inconsistent values in gender, age, and education, otherwise we dropped those observations with missing data.[6] In 2010, the Portuguese Classification of Occupations (CPP) changed in the sense of grouping

[4] Public administration, self-employment and nonmarket services are not covered by QP.

[5] Workers in agricultural and fishery were also excluded from the analysis, as well as workers in the islands of Açores and Madeira, due to their lack of representativeness.

[6] The top and bottom 1 percent observations in wages were excluded from the sample. Multiple job-holder workers were also dropped.

together specific occupations. Whenever possible we recode occupations according to their new classification at the three-digit level, so we can match them with ISCO-08 classification. We also recode the Industry Classification under the ISIC Rev. 3 whenever possible. Thus, our final sample is composed of 31,721 recent graduates (bachelors, graduates, and masters) who entered the private sector in mainland Portugal in 2006 or 2007 aged less than 30 years.

In order to obtain a measure of overqualification, we use the O*NET database successor of labor's Dictionary of Occupational Titles (DOT). The O*NET database contains detailed measures of the importance of tasks that can be recoded into ISCO-08 classification.[7] Acemoglu and Autor (2011) proposed to group these tasks into five categories that go from more demanded, more complex jobs to less demanding jobs given the intensity of their use (see Table 2.1 for details).[8] For each occupation we obtained an index for the intensity of each of the five following categories of tasks: (i) non-routine cognitive analytical task (NR-C.A); (ii) non-routine cognitive interpersonal task (NR-C.I); (iii) routine cognitive task (RC); (iv) routine manual task (RM); (v) non-routine manual physical task (NR-M.P).

*INSERT TABLE 2.1 HERE*

These data were collected using the Standard Occupational Classification (SOC).[9] We recode and normalized the O*NET data into a two-digit ISCO-08 coding by applying a crosswalks between them.[10] Then we use the weighted average of each occupation to obtain the relative importance of tasks within a two-digit occupation code, i.e., all tasks measures are standardized to have a mean of zero and a cross-occupation standard deviation of one.

[7] We are very thankful to Miguel Portela for his generous help on this matter. For further information, please see https://www.onetonline.org.

[8] Acemoglu and Autor (2011) use the importance scales of O*NET database to construct five tasks categories.

[9] The Standard Occupational Classification contains links to major groups, the complete hierarchical structure, broad occupational definitions, and detailed occupational definitions are available at https://www.bls.gov.

[10] Data and codes were prepared by the Institute for Structural Research. For further details please consult www.ibs.org.pl/resources.

### 2.3.2 Measuring Overqualification

The concern raised by Altonji, Kahn, and Speer (2016) toward the choice of the field of study is important in terms of career trajectories. In fact, the skills and tasks that a worker acquires during his/her studies is totally different given the choice of the field of study. Moreover, those skills and tasks are important in terms of career prospect because workers will use them to find a suitable job and to increase their job opportunities. For a given occupation, firms may require more or less complex tasks. For example, some graduate occupations may not make a great use of abstract competences, while others occupations, such as management, need more interpersonal cognitive capacities to lead workers. Suppose that a worker has a specific field of study like economics or engineering. During his/her training, the worker will acquire more analytic tasks that potentially fit the firm's tasks requirement. It is expected that the worker will use them in his/her future career. In turn, other fields of study, such as arts or nursing, require higher interpersonal skills. Theoretically, both previous examples are considered as graduate occupations even though they totally differ in terms of tasks or skills.

Having this in mind, in our analysis, we will consider that an individual is working in a graduate job if he makes a high use of non-routine cognitive analytical tasks, or non-routine cognitive interpersonal tasks.[11] A worker is considered well-matched if the value of the standardized tasks "non-routine cognitive analytical" or "non-routine cognitive interpersonal" is greater than one. Additionally, a worker is considered overqualified if he works in a non graduate job. In other words, if the value of the standardized task "non-routine cognitive analytical" and "non-routine cognitive interpersonal" are simultaneously smaller than one. In Appendix A.1. we present the classification of occupations categories at the two-digit level (ISCO-08) and in Appendix A.2. we provide a summary of the relative importance of tasks within a two-digit occupation code (ISCO-08) according to the five categories proposed by Acemoglu and Autor (2011). In the last column of Appendix A.2. we report the mode of education (in years) for the stock of workers employed in a 2-digit (ISCO-08) occupation. According to our classification, well-matched recent graduates are working in occupations that correspond

[11] Recall that we restrict our analysis to workers having at least a bachelor degree who entered the labor market for the first time in 2006 or 2007.

to major groups 1 and 2 as well as to occupation code 54.[12] Overqualified graduates are working in the remaining major groups. For example, working in occupation with ISCO-08 code 42 "Customer direct support staff" requires a low intensity of non routine cognitive analytic tasks (-0.31), non routine cognitive interpersonal tasks (-0.24), routine manual tasks (-0.42), and non routine manual physical tasks (-0.92) as well as a high intensity of routine cognitive tasks (1.55). As such, a newly graduate that enter the labor market in this occupation is considered to be overqualified inasmuch as the value of "non-routine cognitive: analytical" and "non-routine cognitive: interpersonal" are simultaneously lower than one. Using the mode of education at a 2-digit level, well-matched workers are working in graduate jobs (the mode of education is at least 15 years of schooling), corresponding to major groups 1 and 2, with the exception of occupations 13 and 14.[13] Overeducated workers are working in the remaining major groups (occupations that do not require a graduation and whose mode of education is lower than 15 years).

*INSERT APPENDIXES A.1 AND A.2 HERE*

### 2.3.3 Descriptive Statistics

As a starting point, we provide a statistical comparison between recent graduate workers who entered the labor market overqualified for their job and those who enter the labor market well-matched. According to Table 2.2, 13,709 out of 31,721 recent entrants are overqualified, i.e., 43.2% of the total of recent workers. Recent graduate workers accept a job that do not match their skills because they may want to develop career opportunities or to gain experience (Rubb, 2003). For example, the career mobility theory supports that overskilling serves as a “stepping-stone” to better future jobs, and affects mostly new entrants who lack work-experience.

*INSERT TABLE 2.2 HERE*

[12] Occupation 54 corresponds to "protective and safety services workers", including, for example, police officers or Republican National Guard. According to our classification, it corresponds to a graduate job, therefore workers in this occupation are considered well-matched. However, the mode of education in this occupation is 9 years of schooling, suggesting that a recent graduate working in this occupation is overeducated according to the number of years of schooling.

[13] This result is consistent with previous studies that found that a larger proportion of entrepreneurs/directors in Portugal has lower levels of education (e.g., Rocha *et al.*, 2015).

In our data, newly graduate workers are predominantely women (they are almost as twice as of males), which is an interesting result. Almost 62% of recent graduate workers that are well-matched are women, while 66% of overqualified are also women. The results obtained in most empirical studies of spatial models found that females are more likely to be overeducated in their occupation than men (e.g., Büchel and Battu, 2003; Quintini, 2011).

In Tables 2.3 and 2.4 we report some comparative statistics for newly graduate workers according to their occupations and levels of schoolling. We found that the majority of well-matched workers are employed in major groups 1 and 2; in fact, in these occupations workers make a great use of non-routine cognitive analytical or non-routine cognitive interpersonal skills. We also found that the most frequent occupations among overqualified workers are, for males, occupations categories of major group 3 (51%), and, for females, occupations categories of major group 4 (49%). Regarding the level of education, for both well-matched and overqualified workers, the great majority have a graduation irrespective of the gender. Indeed, 87.54% of male workers that are well-matched have a graduation, whereas 81.35% of overqualified males have the same degree. On the other hand, the great majority of females who are well-matched in their job have a graduation (92%), while 85% of overqualified females have the same level of schooling. There is a larger percentage of bachelors in the group of overqualified workers than in the well-matched group.

*INSERT TABLES 2.3 AND 2.4 HERE*

Table 2.5 shows the percentage of overqualified and well-matched workers according to their field of study. Comparing to well-matched workers, and for both genders, graduates in education, engineering, and medicine are less likely to be overqualified.

*INSERT TABLE 2.5 HERE*

## 2.4 Persistence of Overqualification

### 2.4.1 Empirical Strategy

In this section we analyse the persistence of the overqualification status for the individuals that entered for the first time in the labor market in 2006 or 2007. In particular, we analyse the transitions out of overqualification considering the duration of the first spell of a mismatched job. The duration of the first spell corresponds to the time elapsed until they transit to a state that matches the skills required by the job. We rely on duration models in continuous time, assuming that the duration of the first spell is determined by a parametric proportional hazard (PH) model. We opt to use a parametric model, instead of a semiparametric, because the imposition of a hazard function is seen as the best way to improve the efficiency of estimates (Cleves *et al.*, 2010). The hazard an individual faces is a function of the hazard everyone faces, ($h_i(t)$, the baseline hazard that is a function of $t$ alone), modified by his own relevant characteristics (Cameron and Trivedi, 2005).[14] A functional form for $h_i(t)$ is specified in the PH model. In this study we use a Weibull duration distribution which is at the same time a PH model and an accelerate failure (AFT) time model that forces the hazard function to take a particular shape.[15] The shape of the Weibull model assumes a monotonic hazard, where the hazard function at time $t$ for individual $i$ with covariate vector $X_i$ is:

$$h_i(t) = \exp(\boldsymbol{X}_i\boldsymbol{\beta})pt^{p-1}, \tag{2.1}$$

where $p$, the shape parameter, and regression coefficients $\beta$ are estimated from the data.[16]

In addition, following the existing literature, we control for the effects of individual characteristics (age, gender, nationality, education, field of study, cohort of workers), job characteristics (occupation, type of contract) and firm characteristics (industry, firm size, location) at entry. Including this set of covariates ($\boldsymbol{X}$), the PH model turns to:

$$h(t|X) = pt^{p-1}\lambda \tag{2.2}$$

[14] The hazard rate is the instantaneous probability of leaving a state conditional on survival at time $t$.

[15] As we are more interested in predicting the time to transition, we believe that a parametric assumption would be more appropriate. To choose the best fitting distribution, we use both the Cox-Snell residuals and the Akaike Information Criterion (AIC). The preferred model is the one with the minimum AIC. The Cox-Snell residuals consider the distribution and estimated parameters from the regression. We estimate regression models by maximum likelihood (Exponential, Weibull, Log-normal and Log-logistic). The comparison between the four graphs show that the distribution that fits the data better are, by order of importance, Weibull, Log-normal, Log-logistic and Exponential, which is confirmed by the information criteria.

[16] $p$ determines if the hazard is increasing, decreasing or constant over time.

Regressors are introduced by letting $\lambda$ to depend on the covariates:

$$\lambda_i = e^{\boldsymbol{X}_i \boldsymbol{\beta}} \tag{2.3}$$

And the survivor function is:[17]

$$S(t) = e^{(-\lambda t)^p} \tag{2.4}$$

As mentioned above, we consider two cohorts of newly graduates who were overqualified at the begin of their careers. The first cohort refers to workers who entered into the labor market for the first time in 2006 and a second cohort of those who entered the labor market in 2007. We observe both cohorts from the moment they enter the labor market (2006 or 2007) to the moment of their first transition, when they first become employed in an occupation that matches their skills or until the last period observed in the data (censored observations). During the time spell, they are at risk of making a transition to an adequately matched job. Recall that overqualification refers to individuals who enter the labor market for the first time in 2006 or 2007 in occupations within a 2-digit code making a low use of non-routine cognitive analytical tasks and non-routine cognitive interpersonal tasks (the value of the standardized task is lower than one). In some situations, no transition is made over the analysed period. In this case, these observations are censored (right-censoring). An overqualified spell ends when the worker transits for the first time to a new employment spell that matches the skills required by the job.

Individual differences measured by the regressors (observed heterogeneity) and unmeasured by the regressors (unobserved heterogeneity) both can affect survival time. In our analysis, we control for individual's unobserved heterogeneity, as it can affect the influence of time dependence on the exit rate (e.g., Heckman and Singer, 1984; Jenkins, 2005). In the present context, we will follow the existing literature and use the term frailty to account for unobserved heterogeneity (Vaupel *et al.* 1979; Lancaster, 1979; Gutierrez, 2002).[18] Overdispersion is caused either by misspecification or omitted covariates. When unobserved heterogeneity is ignored, its impact is counfounded with

[17] The survivor function represents the probability that duration equals or exceeds $t$.

[18] A frailty model is a generalization of a survival regression model (Gutierrez, 2002).

that of the baseline hazard. In duration models, the introduction of unobserved heterogeneity leads to mixture models. When modelling individual heterogenity, a frailty model is just the standard parametric model with the addition of a new parameter (the variance of the unobserved observation-specific effect) and a new definition of the survivor function. For the choice of the heterogeneity function, in continuous time, the distributions that are mostly used are Gamma and Inverse Gaussian distributions (Jenkins, 2005). In the present study, we assume an Inverse Gaussian distribution.[19]

A frailty model introduces an unobservable multiplicative effect, $\alpha$, on the hazard, so that for the $j$th observation in the $i$th group, a frailty model writes as:

$$h(t_{ij}|\alpha_{ij}) = \alpha_{ij} h(t_{ij}) \tag{2.5}$$

where $\alpha$ is assumed to have mean one and variance $\theta$. If workers have $\alpha > 1$ they will be more at risk to transit, i.e., to become adequately qualified; conversely, if $\alpha < 1$, workers will survive longer in the state of overqualified. Suppressing the index, equation (5) becomes:

$$h(t|\alpha) = \alpha h(t) \tag{2.6}$$

Since $\alpha$ is not observable, it has to be integrated out of $S(t|\alpha)$ to obtain the unconditional survival function (Gutierrez, 2002). The survivor function $S(t)$ that corresponds to $h(t)$ given the frailty is as follows:

$$S(t|\alpha) = exp\left\{-\int_0^t h(u|\alpha)du\right\} = \exp\left\{-\alpha\int_0^t \frac{f(u)}{S(u)}du\right\} = \{S(t)\}^{\alpha} \tag{2.7}$$

Let $g(\alpha)$ be the probability density function of $\alpha$, the inverse-Gaussian distribution has density:

$$g(\alpha) = \left(\frac{1}{2\pi\theta\alpha^3}\right)^{1/2} exp\left\{-\frac{1}{2\theta}\left(\alpha - 2 + \frac{1}{\alpha}\right)\right\} \tag{2.8}$$

[19] See Hougaard (1984) for a discussion on the distribution of the frailty for a gamma and an inverse Gaussian distributions.

With frailty models, the unconditional probability of survival is described by $S_\theta(t)$ instead of $S(t)$. Therefore, when $\alpha$ follows an Inverse Gaussian distribution, the survival function turns to:

$$S_\theta(t) = exp\left\{\frac{1}{\theta}(1 - [1 - 2\theta ln\{S(t)\}]^{1/2}\right\} \tag{2.9}$$

The frailty model reduces to $S(t)$ when there is no heterogeneity. A worker will have probability of survival past time $t$ equal to $\{S(t)\}^\alpha$, whereas $S_\theta(t)$ is the measure of the proportion of workers that survive past time $t$.

### 2.4.2 Empirical Results

For the purpose of duration analysis we organized the data in time-span data. In our sample we have 13,709 total observations at risk of transition. Of this total, 3,963 individuals moved from an overqualified job to a well-matched job, while the remaining 9,746 individuals remained in the initial status in the period under analysis. Some individuals may have left *QP* without having transited for several reasons: they may have become unemployed, they may have moved to the public sector, they may have become self-employed or they are in maternity leave.

Nonparametric models are very usefull for descriptive purposes as they allow us to know the shape of the raw (unconditional) hazard or survival function.[20] Figure 2.1 plots the survival rates, i.e., the lenght of time that workers remain overskilled, using the Kaplan Meier estimator (Kalbfleish and Prentice, 2002).

*INSERT FIGURE 2.1 HERE*

Consistent with the previous literature, this figure seems to confirm that overqualification is a permanent mismatch where newly graduates workers get trap. As reported by Table 2.6 (last row of column 4), at the end of the time span, 53.6% of newly graduates overqualified at entry fail to move to an adequate jobs and remain overqualified in 2012, showing a strong entrapment effect of overqualification. This finding is in line with other studies who concluded that many invidKuals still remain overqualified

[20] Using nonparametric models, we let the data speak for itself, no assumptions are made about either the distribution of failure times or how covariates change the survival time.

for long periods of time. Robst (1995), using data for the US, found that almost 60% of overeducated workers in 1976 were still overeducated nine years later. According to Kiersztyn (2013), based on data from the Polish Panel Survey (POLPAN), the youngest cohorts, workers aged from 26 to 35 years old in 2008, face a higher risk of persistent overqualification than other cohorts during a crisis. Frenette (2004), using data from the Canadian National Graduates Surveys (NGS), find that more than half of all graduates of master's programs are still overqualified five years after graduating. In a recent paper, Meroni and Vera-Toscano (2017), using the 2005 REFLEX data restricted to the European countries and Norway, find evidence that overeducation at the beginning of a career leads to a greater likelihood of being overeducated on the future, while Erdsiek (2017), based on data for Germany, refers a highly persistence of overqualification among tertiary graduates over the first ten years of their career cycle.

*INSERT TABLE 2.6 HERE*

In order to analyse the impact of conditions at entry in the hazard rate, we estimated a parametric continuous time model, with time-invariant covariates, as we are interested in analysing how covariates at the time of entry affect the survival time. We estimated a PH model where the baseline duration follows a Weibull distribution. We start by describing the set of included time-invariant covariates in our regressions. All our regressions include worker, job and firm specific variables. The descriptive statistics of the variables included in the PH model are presented in Table 2.7.

*INSERT TABLE 2.7 HERE*

In Table 2.8, we present the results obtained from the estimation of the continuous time hazard model with and without controlling for unobserved heterogeneity as described above.

*INSERT TABLE 2.8 HERE*

In the first column we report a standard Weibull duration model that disregards unobserved heterogeneity. In the second column we report the estimates of the coefficients obtained from the frailty model using the same covariates, that controls for unobserved heterogeneity. In the PH models, the sign of the coefficients indicates how a covariate affects the harzard rate.[21] A positive coefficient increases the hazard rate of transition, which means that it reduces the expected duration of the mismatch. On the opposite, a negative coefficient decreases the hazard rate of transition.

The estimated parameter $p$ is greater than 1, meaning that the hazard is monotonically increasing with time. Thus, the hazard of transition increases over time as spells in overqualification grow larger, implying significant positive duration dependence.

Looking at the first estimation, we can see that more educated workers, males, natives, those whose field of study is "engineering and technology" and are located in an urban location face higher hazard rates and hence transit more quickly to an adequate job. On the contrary, females and foreigners present a negative coefficient, indicating that they transit less quickly, as such, they survive longer overqualified. Overall, our results suggest that newly graduate males transit earlier to a job for which they are adequately qualified, conversely, newly graduate females persist longer in a skill mismatch.

Looking at the frailty model, it turns out that frailty is statistically significant. As such, the frailty model is preferred to the reference non frailty model according to the relevant likelihood ratio test.[22] Furthermore, we found that the estimated parameter $p$ in the frailty model is larger than in the reference model, i.e, the baseline slopes upwards to a greater extent.

Once we control for frailty, we found that the magnitude of the regression coefficients are larger than in the PH model, indicating that workers transit even sooner out of overqualification. In other words, they survive even less time overqualified when controlling for unobserved heterogeneity. Overall, our results suggest that overqualified workers that remain overqualified at high durations tend to have lower hazard rates. This can be explained by the fact that, as time goes by, the population is more and more crowded by "slow-type" workers and not because long time spending in an overqualified job reduces the probability to exit overeducation (Clark *et al.*, 2017).

[21] The hazard function is defined as the probability that the spell is completed at $t$, given that it has not been completed before $t$. The value of the hazard function is called the hazard rate. The hazard ratio (failure/survival) is equal to $\exp(\hat{\beta})$, if this ratio is smaller than one, it implies a negative effect on the hazard.

[22] We reject the null that $\theta = 0$ suggesting that unobserved heterogeneity is affecting our model. The LR test compares the Weibull frailty model to the reference Weibull.

## 2.5 Overqualification and Subsequent Wages

### 2.5.1 Estimation Methodology

The aim of this Section is to compare the subsequent wage growth of recently graduates who have entered the labor market overqualified with the wage growth of recently graduates who have entered the labor market employed in a well-matched job. The data used in this section comprise 56,475 workers'observations who entered the labor market overqualified and 75,826 workers observations who entered well-matched. All together, we identify 132,301 workers'observations who entered the labor market for the first time in 2006 or 2007.

According to subsequent transitions, and for comparison purposes, we divided the group of overqualified workers at entry into two sub-groups:

(i) those who never transited to a well-matched job;

(ii) those who transited at least once for a well-matched job.

Regarding the control group, and to check the robustness of our results to different definition of the control group, we also splitted the latter into two sub-groups:

(i) those who remained continuously employed in a well-matched job;

(ii) those who transited at least once for a mismatched job.

Table 2.9 summarizes the definitions of the treated and control groups namely, recent graduates who were overqualified at the begin of their careers and never transited to a well-matched job (**continuously overqualified**); recent graduates who were overqualified at their entry in the labor market and transit at least once for a well-matched job (**initial overqualified**); recent graduates who entered the labor market well-matched and never transited (**continuously well-matched**), and recent graduates well-matched at their entry who transited to a mismatched job (**initial well-matched**) at least once.

*INSERT TABLE 2.9 HERE*

Table 2.10 presents the number of observations in each period for the four groups aforementioned under the analysed period.

*INSERT TABLE 2.10 HERE*

Table 2.11 reports the real average hourly wages (in logs) for each group. The hourly wages correspond to total regular payroll (base wage and regular payments) over normal hours worked in the reference month converted into 2010 constant prices using the Consumer Price Index (CPI). Regarding wages, and unconditionally, we observe significative differences between the four groups of workers. On average, the wages are higher for the group of continuously well-matched and initial well-matched workers when compared to workers who entered the labor market overqualified.

*INSERT TABLE 2.11 HERE*

Figure 2.2 shows the evolution of wages for the four groups of workers considered, without controlling for any workers' characteristics. In Figure 2.2 we can see that at the beginning of their career, workers who start their careers well-matched earn more than workers who entered overqualified. Even though the latter seem to exhibit a higher wage growth in the first years of their careers, in particular those that moved to a well-matched job. In any case, in the end of the analysed period their wages remain below the wages of workers who entered well-matched.

*INSERT FIGURE 2.2 HERE*

Comparing the evolution of wages by gender, Figure 2.3 shows that overqualified females receive a lower wage when they start their career when compared to overqualified males. At the end of the analysed period, they still receive less than males. On the opposite, female workers that enter the labor market well-matched receive almost the same wages as well-matched males. However, the former exhibit a lower wage growth, implying that, by the end of the analysed period, well-matched females receive less than well-matched males. Furthermore, the wage profile of initial overqualified workers and initial well-matched workers is very identical in the first five years after entry.

*INSERT FIGURE 2.3 HERE*

We further examine if being overqualified at the beginning of a career is associated with future lower wages when controlling for worker, firm, and job characteristics. Our empirical model includes controls for both worker's and firm's unobserved permanent heterogeneity. In fact, there is evidence that overqualification may work as a mechanism to compensate the lack of worker's experience or ability. Thus, not controlling for unobserved individual heterogeneity can over or under-estimate the effect of skill mismatch on wages. On the other hand, workers are not randomly affected to occupations, i.e., they can self-select into occupations according to their comparative advantages (e.g., Cortes, 2016). Groes, Kircher, and Manovskii (2015) also found that more able (less able) workers are more likely to switch to occupations with higher (lower) average wages. Furthermore, less able workers are more likely to be overskilled, because they may not be able to find a job that matches their skills (Chevalier, 2003). Employers may accept to hire overqualified workers, temporarily, to verify if they are prepared for more demanding or more skilled positions. Thus, we opt to control for both firm permanent unobserved heterogeneity and workers unobserved heterogeneity and to estimate a two high-dimensional fixed-effects wage equation, following the procedure described in Guimarães and Portugal (2010).

The full model writes as:

$$\ln w_{it} = \gamma OQ_i + \lambda_1 YSE_{it} + \lambda_2 YSE_{it}^2 + \delta_1(OQ_i \times YSE_{it}) + \delta_2(OQ_i \times YSE_{it}^2) + \boldsymbol{\beta} X_{it} + \alpha_i + \theta_j + \varepsilon_{it} \tag{2.10}$$

where $w_{it}$ is the log of hourly wages (in real euros) for each individual $i$ at firm $j$ in year $t$. $OQ$ is a dummy variable taking the value one if a recent graduate worker $i$ entered the labor market overqualified. Thus, this variable is equal to one for continuously overqualified and initial overqualified workers - the treated group - and equals to zero for the control group (continuously well-matched and initial well-matched workers). In order to compare the behavior of wages for both groups of workers over the analysed period, we included in the regression a quadratic trend for the years since entry in the labor market ($YSE$) and interaction terms between the dummy $OQ$ and years since entry ($OQ \times YSE$) and its square ($OQ \times YSE^2$). $X_{it}$ is a set of control variables for worker and job characteristics, $\alpha_i$ is a worker fixed effect, $\theta j$ corresponds to the firm fixed effect, and $\varepsilon_{it}$ is a random error term assumed to be uncorrelated with the regressors. The descriptive statistics of the variables included in the wage model are reported in Table 2.12.

*INSERT TABLE 2.12 HERE*

### 2.5.2 Empirical Results

Table 2.13 reports the estimates of the wage model defined in equation (10). For comparison purposes, two additional regressions were reported. The first regression, reported in column (1), presents estimates from the standard OLS model. The second model, in column (2), accounts for worker fixed-effects whereas the third model, in column (3), reports estimates when controlling for both worker and firm fixed-effects.[23]

*INSERT TABLE 2.13 HERE*

The OLS estimates indicate that, ceteris paribus, overqualified workers at entry face a wage penalty of 15%.[24] The OLS estimates indicate that overqualified workers experience a higher wage growth over the analysed period when compared with their similar counterparts well-matched.[25] However, this advantage tends to disappear once we control for both worker and firm unobserved heterogeneity, suggesting that overqualified workers experience a similar pattern in terms of wage growth when compared with the control group. Most of the existing studies have been suggesting that overeducated workers exhibit lower or similar wage growth rates than well-matched workers (e.g., Buchel and Mertens, 2004; Korpi and Tåhlin, 2009).[26]

Finally, in order to test to what extent the wage growth of overqualified workers that were able to move to a well-matched job differ from the wage growth of continuously overqualified workers, we re-estimate the wage equation including interaction terms between a dummy variable that takes the value one for continuously overqualified workers (0 otherwise) and the quadratic term in YSE and a dummy variable that takes the value one for initial overqualified workers (0 otherwise) and the quadratic term in YSE (see Table 2.14). This robustness check seeks to provide a better understanding of the evolution of the wages between continuously overqualified workers (never transited) and workers that were overqualified at entry. We find that the wage development of initial overqualified workers is statistically different from the control

[23] The two-way high-dimensional fixed effects models were estimated using the "reghdfe" stata command developed by Correia (2016).

[24] The exact value is computed $(\exp(\hat{\beta}) - 1) * 100$.

[25] The coefficients of the control variables have the expected sign. Males, foreigners, more educated and older graduates, and workers who hold a permanent contract earn higher wages than their similar counterparts. Regarding the field of study, graduates in medicine and social science have, on average, higher wages than similar graduates in engineering and technology.

[26] As robustness checks we have re-estimated the three models reported in Table 2.13 using different definitions of the control and treated groups. In Appendix A.3 we excluded from the control group the initial well-matched, i.e., the well-matched workers at entry that later on switched to a mismatched job. In Appendix A.4 we excluded from the treated and control groups workers that changed status over the analysed period, i.e., the initial well matched and the initial overqualified. Overall, the results remain qualitatively identical to the ones reported in Table 2.13.

group and the treated group of continuously overqualified workers, exhibiting, on average, a higher wage growth. In other words, five years since labor market entry the rate of wage growth of initial overqualified workers that moved to a well-matched job exceeds, by about 12 p.p., the rate of wage growth of workers who entered the labor market well-matched.

*INSERT TABLE 2.14 HERE*

Finally, figure 2.4 displays the empirical distribution of permanent worker observed and unobserved heterogeneity for the population of continuously overqualified, initial overqualified, continuously well-matched and initial well-matched workers. Worker permanent heterogeneity is proxied by the estimates of the worker fixed effect filtered from firm permanent observed and unobserved heterogeneity obtained by the estimation of the full-model. The graph shows that the empirical distributions of the worker fixed effect for well-matched workers are more shifted to the right, while for the groups of overqualified workers are more shifted to the left, in particular for continuously overqualified workers. In other words, well-matched workers seem to correspond to a higher-ability group of workers, while continuously overqualified workers seem to be the lowest-ability group of employees.

*INSERT FIGURE 2.4 HERE*

## 2.6 Concluding Remarks

The aim of this paper is to analyse the career dynamics of overqualified workers in terms of future employment and wage prospects using a sample of newly graduates in Portugal. In particular, the purpose of this essay is to provide a better understanding of the impacts of this phenomenon in the medium-long-run on subsequent employment and wage prospects. Exploring a large matched employer-employee data set over the 2006-2012 period, we find that overqualification is a permanent phenomenon for a great majority of workers: six years after entering the labor market, 63% of the workers that entered overqualified remain mismatched. The parametric analysis show that more educated workers, natives and males are more likely to transit into out of overqualification, whereas females, foreigners, and less educated workers persit longer in a skill mismatch. Once we account for unobserved heterogeneity, we find that the hazard rate out of overqualification increases. This result is also corroborated by the estimates of the worker fixed-effect that indicate that the well-matched seem to correspond to a higher ability group of employees, while the continuously overeducated seem to correspond to a lower ability group of workers.

Regarding wages, the results reveal that at entry, and unconditionally, overqualified workers earn lower wages when compared with well-matched workers, but this gap tends to diminish for overqualified workers that were able to move to a well-matched job over the analysed period. Taking into account workers and firms observed and unobserved permanent heterogeneity, the estimates indicate that five years after entering the labor market, overqualified individuals that were able to switch to a well-matched job experience a wage growth that exceeds the wage growth of their similar well-matched counterparts in 12 percentage points.

# Bibliography


[1] Acemoglu, D., Autor, D. (2011). "Skills, tasks and technologies: Implications for employment and earnings". Elsevier, Part B of Handbook of Labor Economics, 4: 1043-1171.

[2] Alba-Ramirez, A. (1993). "Mismatch in the Spanish labor market: Overeducation?". Journal of Human Resources, 27(2): 259-278.

[3] Altonji, JG., Kahn, L.B., Speer, J.D. (2016). "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success". Journal of Labor Economics, 34(1): 361-401.

[4] Baert, S., Cock, B., Verhaest, D. (2013). "Overeducation at the start of the career: Stepping stone or trap?" Labour Economics, 25: 123-140.

[5] Becker, G. S. (1964). "Human capital". New York: National Bureau of Economic Research (NBER).

[6] Blázquez, M., Budría, S. (2012). "Overeducation dynamics and personality". Education Economics, 20(3): 260-283.

[7] Boll, C., Leppin, J. S., Schömann, K. (2016). "Who is overeducated and why? Probit and dynamic mixed multinomial logit analyses of vertical mismatch in East and West Germany". Education Economics, 24(6): 639–662.

[8] Büchel, F. (2002). "The effects of overeducation on productivity in Germany – The firms' viewpoint". Economics of Education Review, 21(3): 263-275.

[9] Büchel, F. and Battu, H. (2003). "The theory of differential overqualification: does it work". Scottish Journal of Political Economy, 50(1): 1-16.

[10] Büchel, F., Mertens, A. (2004). "Overeducation, undereducation, and the theory of career mobility". Applied Economics, 36(8): 803-816.

[11] Cameron, A. C., Trivedi, P. K. (2010). "Microeconometrics using Stata". Revised Edition, Stata Press.

[12] Chevalier, A. (2003). "Measuring Over-education". Economica, 70(279): 509-531.

[13] Clark, B., Joubert, C., Maurel, A. (2017). "The career prospects of overeducated Americans". IZA Journal of Labor Economics, 6(3).

[14] Cleves, M., Gould, W., Gutierrez, R. G., Marchenko, Y.V. (2010). "An Introduction to Survival Analysis Using Stata". 3rd ed. College Station, Tex: Stata Press.

[15] Correia, S. (2016). "REGHDFE: Stata Module to Perform Linear or Instrumental-Variable Regression Absorbing Any Number of High-Dimensional Fixed Effects". Statistical Software Components, Boston College, Department of Economics.

[16] Cortes, G. M. (2016). "Where have the middle-wage workers gone? a study of polarization using panel data". Journal of Labor Economics, 34(1): 63–105.

[17] Dolton, P., Vignoles, A. (2000). "The incidence and effects of overeducation in the U.K. graduate labour market". Economics of Education Review, 19(2): 179-198.

[18] Dorn, D., Sousa-Poza, A. (2005), "Over-qualification: Permanent or transitory", mimeo, University of St Gallen, Switzerland.

[19] Erdsiek, D . (2017). "Dynamics of overqualification: Evidence from the early career of graduates". ZEW Discussion Papers, N°17-020.

[20] Freeman, R. (1976). "The overeducated American". Academic Press.

[21] Frei, C., Sousa-Poza, A. (2012). "Overqualification: Permanent or transitory?". Applied Economics, 44(14): 1837-1847.

[22] Frenette, M. (2004). "The overqualified Canadian graduate: The role of the academic program in the incidence, persistence, and economic returns to overqualification". Economics of Education Review, 23(1): 29-45.

[23] Groes, F., Kircher, P., Manovskii, I. (2015). "The U-shapes of occupational mobility". Review of Economic Studies, 82 (2): 659-692.

[24] Groot, W., Maassen van den Brink, H. (2000). "Overeducation in the labor market: a meta-analysis". Economics of Education Review, 19: 149-158.

[25] Guimarães. P., Portugal, P. (2010). "A Simple Feasible Alternative Procedure to Fit Models with High-Dimensional Fixed Effects". The Stata Journal, 10(4): 628-649.

[26] Gutierrez, R. G. (2002). "Parametric frailty and shared frailty survival models". The Stata Journal, 2(1): 22-44.

[27] Heckman, J., Singer, B. (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data". Econometrica, 52(2): 271-320

[28] Hougaard, P. (1984). "Life Table Methods for Heterogenous Populations: Distributions Describing the Heterogeneity". Biometrika, 71: 75-83.

[29] Jenkins, S. P. (2005). "Survival analysis". Unpublished manuscript.

[30] Kalbfleisch, J. D., Prentice, R. L. (2002). "The Statistical Analysis of Failure Time Data, second edition". John Wiley & Sons, Inc, New Jersey.

[31] Kiersztyn, A. (2013). "Stuck in a mismatch? The persistence of overeducation during twenty years of the post-communist transition in Poland". Economics of Education Review, 32: 78-91.

[32] Korpi, T., Tåhlin, M. (2009). "Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000". Labour Economics, 16(2): 183-193.

[33] Lancaster, T. (1979). "Econometric Methods for the Duration of Unemployment". Econometrica 47(4): 939-956.

[34] Mavromaras K., S. McGuinness. (2012). "Overskilling dynamics and education pathways". Economics of Education Review, 31(5): 619-628.

[35] McGuinness, S. (2006). "Overeducation in the labour market". Journal of Economic Surveys, 20(3): 387–418.

[36] McGuinness, S., Pouliakas, K., Redmond, P. (2017). "How useful is the concept of skills mismacth?". IZA DP N. 10786.

[37] Meroni, E. C., Vera-Toscano, E. (2017). "The persistence of overeducation among recent graduates". Labour Economics, 48: 120-143.

[38] Mincer, A. (1974). "Schooling, experience and earnings". NewYork: NBER Press.

[39] Quintini, G. (2011). "Over-qualified or under-skilled: A review of existing literature". OECD Social, Employment and Migration WP No 121. DELSA/ELSA/WD/SEM(2011)6.

[40] Rocha, V., Carneiro, A., Varum Amorim, C. (2015). "Entry and exit dynamics of nascent business owners". Small Business Economics, 45(1): 63-84.

[41] Robst, J. (1995). "College quality and overeducation". Economics of Education Review, 14(2): 221–228.

[42] Rubb, S. (2003). "Overeducation: a short or long run phenomenon for individuals?". Economics of Education Review, 22: 389–394.

[43] Sattinger, M. (1993). "Assignment models of the distribution of earnings". Journal of Economic Literature, 31(2): 831–880.

[44] Sicherman, N. (1991). "Overeducation in the labor market". Journal of Labor Economics. 9(2): 101–122.

[45] Sicherman, N., Galor, O. (1990). "A Theory of Career Mobility". Journal of Political Economy, 98(1): 169-192.

[46] Sloane, P. J. (2003). "Much ado about nothing? What does the over-education literature really tell us?" In F. Buchel, A. de Grip, & A. Mertens (Eds.), Overeducation in Europe. Current Issues in Theory and Policy: 11–45. Cheltenham: Edward Elgar.

[47] Thurow, L., C. (1975). "Generating Inequality". New York: Basic Books.

[48] Tsang, M. C., Levin, H. M. (1985). "The Economics of Overeducation ". Economics of Education Review, 4(2): 93-104.

[49] Vaupel, J. W., Manton, K.; Stallard, E. (1979)."The impact of heterogeneity in individual frailty on the dynamics of mortality". Demography,16: 439-454.

[50] Verhaest, D., Schatteman, T. (2010). "Overeducation in the early career: an analysis using sequence techniques". (HUB Research Papers in Economics and Management, 2010/09).

[51] Verhaest, D., Van der Velden, R. (2013). "Cross-country differences in graduate overeducation". European Sociological Review 29(3): 642-653.

# Tables and Figures

Table 2.1 – O*NET Data descriptors (Acemoglu and Autor, 2011)

| Classification | Tasks |
|---|---|
| Non-routine Cognitive Analytical | Analyzing Data or Information<br>Thinking Creatively<br>Interpreting the Meaning of Information for Others |
| Non-routine Cognitive Interpersonal | Coaching and Developing Others<br>Guiding, Directing, and Motivating Subordinates<br>Establishing and Maintaining Interpersonal Relationships |
| Routine Cognitive | Importance of Being Exact or Accurate<br>Importance of Repeating Same Tasks<br>Structured versus Unstructured Work |
| Routine Manual | Pace Determined by Speed of Equipment<br>Spend Time Making Repetitive Motions<br>Controlling Machines and Processes |
| Non-routine Manual physical | Spatial Orientation<br>Manual Dexterity<br>Operating Vehicles, Mechanized Devices, or Equipment<br>Spend Time Using Your Hands to Handle,<br>Control, or Feel Objects, Tools, or Controls |

Table 2.2 – Distribution of overqualified and well-matched workers by gender (Portugal, 2006-2012)

| | Overqualified | | | Well-matched | | |
|---|---|---|---|---|---|---|
| | Males | Females | Total | Males | Females | Total |
| Cohort 2006 | 2135 | 4297 | 6432 | 3385 | 5805 | 9190 |
| Cohort 2007 | 2498 | 4779 | 7277 | 3593 | 5229 | 8822 |
| Total | 4633 | 9076 | 13709 | 6978 | 11034 | 18012 |

Table 2.3 – Distribution of overqualified and well-matched workers at entry by occupations at 2-digit (Portugal, 2006-2012)

| | Overqualified (%) | | Well-matched (%) | |
|---|---|---|---|---|
| Occupations 2 digit-codes (ISCO-08) | Males | Females | Males | Females |
| Legislative power (11, 12, 13, 14) | | | 4.12 | 3.19 |
| Physical science (21, 22, 23, 24, 25, 26) | | | 95.46 | 96.59 |
| Protective and safety services workers (54) | | | 0.42 | 0.22 |
| Science and engineering (31, 32, 33, 34, 35) | 51.05 | 32.28 | | |
| Administrative staff (41, 42, 43, 44) | 36.15 | 48.78 | | |
| Services and sales (51, 52, 53) | 6.34 | 15.67 | | |
| Skilled construction, industry sector (71, 72, 73, 74, 75) | 3.08 | 1.10 | | |
| Plant, machine operators, assemblers (81, 82, 83) | 1.28 | 0.52 | | |
| Unskilled workers (91, 92, 93, 94, 95, 96) | 2.10 | 1.65 | | |
| Total | 100% | 100% | 100% | 100% |

Table 2.4 – Distribution of overqualified and well-matched workers at entry by level of schooling (Portugal, 2006-2012)

| | Overqualified (%) | | Well-matched (%) | |
|---|---|---|---|---|
| Education level | Males | Females | Males | Females |
| Bachelors | 14.88 | 11.36 | 8.55 | 5.20 |
| Graduates | 81.35 | 85.42 | 87.54 | 91.86 |
| Masters | 3.77 | 3.22 | 3.92 | 2.94 |
| Total | 100% | 100% | 100% | 100% |

Table 2.5 – Distribution of overqualified and well-matched workers at entry by field of study (Portugal, 2006-2012)

| | Overqualified (%) | | Well-matched (%) | |
|---|---|---|---|---|
| Field of study | Males | Females | Males | Females |
| Humanities | | | | |
| History, philosophy, etc | 2.46 | 6.36 | 1.26 | 2.94 |
| Arts | 2.82 | 2.80 | 1.70 | 1.45 |
| Professions and applied sciences | | | | |
| Business Science | 21.34 | 20.85 | 11.45 | 9.19 |
| Education | 1.88 | 4.44 | 3.59 | 13.51 |
| Engineering and technology | 21.10 | 6.80 | 33.49 | 7.35 |
| Journalism, media studies, and communication | 1.68 | 2.71 | 1.41 | 2.48 |
| Medicine | 3.65 | 5.90 | 9.25 | 24.82 |
| Social sciences | 6.88 | 10.70 | 5.27 | 9.48 |
| Ignored | 38.19 | 39.44 | 32.58 | 28.78 |
| Total | 100% | 100% | 100% | 100% |

Figure 2.1 – Kaplan-Meier Survival rates

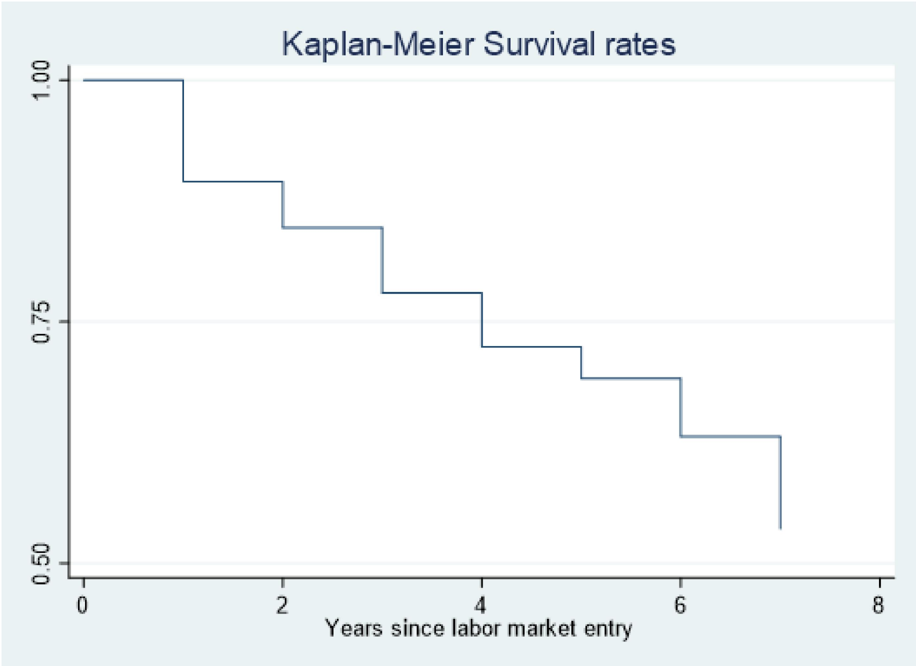

Table 2.6 – Survival rates of overqualified (Portugal, 2006-2012)

| Years since entry | Number of individuals at risk | Number of individuals that transit | Number of censored observations | Survival rates |
|---|---|---|---|---|
| 1 | 13,709 | 1,435 | 485 | 0.895 |
| 2 | 11,789 | 629 | 1,955 | 0.848 |
| 3 | 9,205 | 737 | 1,652 | 0.780 |
| 4 | 6,816 | 487 | 1,350 | 0.724 |
| 5 | 4,979 | 226 | 1,293 | 0.691 |
| 6 | 3,460 | 300 | 2,176 | 0.631 |
| 7 | 984 | 149 | 835 | 0.536 |
| Total | | 3,963 | 9,746 | |

Figure 2.2 – Evolution of hourly wages by match status at entry

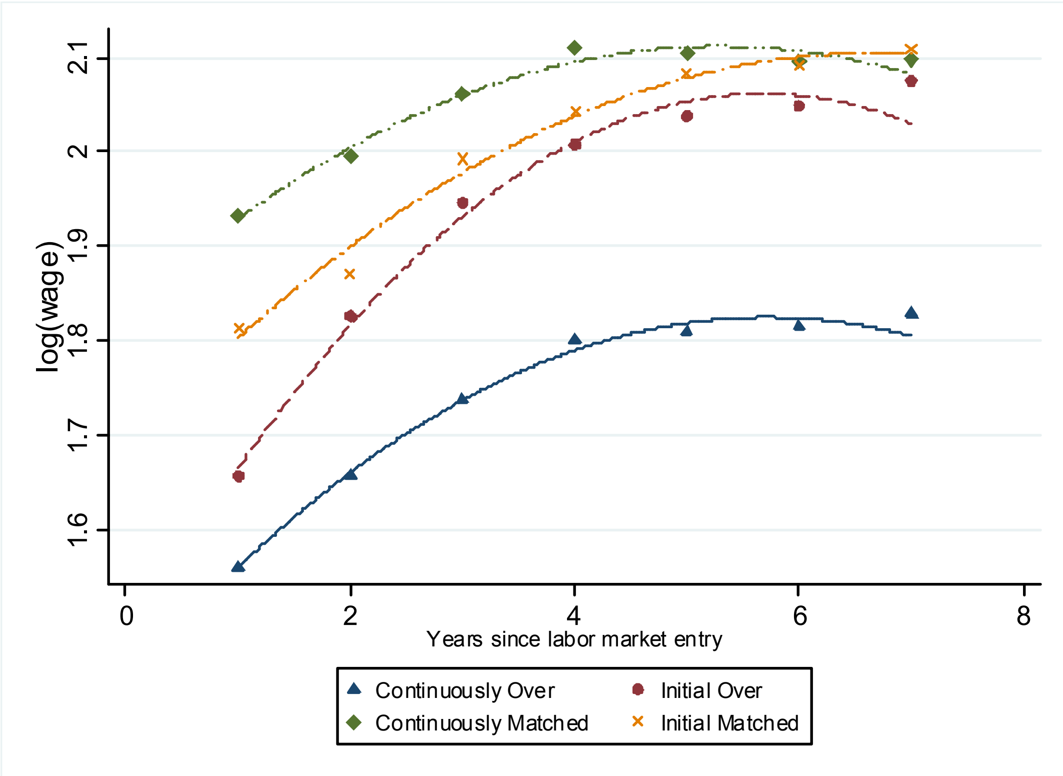

Table 2.7 – Description of variables included in the PH model

| Variables | Description of variables |
|---|---|
| **Individual-level characteristics** | |
| Female | = 1 for females, 0 otherwise |
| Age25-29 | = 1 if greater than 25 years old, 0 otherwise |
| Foreigner | = 1 foreigner, 0 otherwise |
| Education Level: | |
| Bachelor | = 1 if the worker has a bachelor degree, 0 otherwise |
| Master | = 1 if the worker has a master degree, 0 otherwise |
| Field of study: | |
| History, philosophy, etc. | = 1 if the field of study is History, philosophy etc., 0 otherwise |
| Arts | = 1 if the field of study is Arts, 0 otherwise |
| Business Science | = 1 if the field of study is Business, 0 otherwise |
| Education | = 1 if the field of study is Education, 0 otherwise |
| Journalism, media studies, and communication | = 1 if the field of study is Journalism, 0 otherwise |
| Medicine | = 1 if the field of study is Medecine, 0 otherwise |
| Social Sciences | = 1 if the field of study is Social sciences, 0 otherwis |
| Non-defined | = 1 if the field of study is ignored, 0 otherwise |
| Cohort 2007 | = 1 if entered in the labor market for the first time in 2007, 0 otherwise |
| **Job Characteristics** | |
| Type of contract: | |
| Fixed term contract | = 1 if fixed term contract, 0 otherwise |
| Non-defined | = 1 if type of contract is ignored, 0 otherwise |
| Occupation categories 2-digit (ISCO08): | |
| Administrative staff | = 1 if occup. is 41, 42, 43 or 44; 0 otherwise |
| Service and sales | = 1 if occup. is 51, 52, 53 or 54; 0 otherwise |
| Skilled construction, industry sector | = 1 if occup. is 71, 72, 73, 74 or 75; 0 otherwise |
| Plant, machine operators, assemblers | = 1 if occup. is 81, 82 or 83; 0 otherwise |
| Unskilled workers | = 1 if occup. is 91, 92, 93, 94, 95 or 96; 0 otherwise |
| **Firm characteristics** | |
| Firm size | number of employees in the firm (in logs) |
| Urban location | = 1 if the firm is located in an urban area (districts of Porto or Lisbon); 0 otherwise |
| Industry dummies | Dummy variables for each 2-digit industry |

Table 2.8 – Duration models for the first spell in the status of overqualified, Weibull PH regression

| | PH model | Frailty model |
|---|---|---|
| | (1) | (2) |
| Female | -0.2485*** | -0.4291*** |
| | (0.0343) | (0.0618) |
| Age25-29 | 0.0275 | 0.0107 |
| | (0.0342) | (0.0597) |
| Foreigner | −0.3925 | −0.6558** |
| | (0.1346) | (0.2226) |
| Education level (omitted category - bachelor): | | |
| Graduate | 0.2039*** | 0.3511*** |
| | (0.0471) | (0.0809) |
| Master | 0.3765*** | 0.6656** |
| | (0.1111) | (0.1927) |
| Field of study (omitted cat. - engineering & technology): | | |
| History, philosophy, etc. | −0.4013*** | −0.6599*** |
| | (0.0936) | (0.1556) |
| Arts | −0.3127** | −0.5606** |
| | (0.1058) | (0.1857) |
| Business Science | −0.2612*** | −0.4640*** |
| | (0.1254) | (0.0851) |
| Education | −0.0209 | −0.0153 |
| | (0.0993) | (0.1071) |
| Journalism, media studies & communication | −0.3841** | −0.7115** |
| | (0.1140) | (0.2103) |
| Medicine | −0.4057*** | −0.6437*** |
| | (0.1058) | (0.1562) |
| Social Sciences | −0.0557 | −0.0967 |
| | (0.0585) | (0.1909) |
| Cohort 2007 | 0.0297 | 0.0280 |
| | (0.0344) | (0.0569) |
| Fixed term contract | −0.0010 | 0.0197 |
| | (0.0389) | (0.0695) |
| Firm size | −0.0126 | −0.0235 |
| | (0.0084) | (0.0148) |
| Urban location | 0.0794* | 0.1332* |
| | (0.0419) | (0.0737) |
| Constant | −2.7262*** | −2.1547*** |
| | (0.1008) | (0.1906) |
| $p$ | 1.4215 | 2.4534 |
| | (0.0154) | (0.0521) |
| $\theta$ | | 27.3755 |
| Log pseudolikelihood | −10300.94 | −10177.94 |
| $N$ | 13,709 | 13,709 |

Notes: (i) all regressions include a set of occupation and industry dummies (2-digit level), non-defined field of study and contract (ii) worker-cluster robust standard errors in parentheses; (iii) *, **, and *** denote significant at 10%, 5% and 1%, respectively;(iv) the LR test to test $\theta = 0$ is 246,01 allowing us to reject the null hypothesis,as such, unobserved heterogenity is relevant.

Table 2.9 – Definition of the treated and control groups based on match status in first job and subsequent transitions

| Description | Switched job status |
|---|---|
| Overqualified (OQ) | |
| Continuously overqualified | No |
| Initial overqualified | Yes |
| | |
| Well-matched (WM) | |
| Continuously well-matched | No |
| Initial well-matched | Yes |

Table 2.10 – Number of observations by period for treated and control groups

| | Continuously overqualified | Initial overqualified | Continuously well-matched | Initial well-matched |
|---|---|---|---|---|
| $t$ | 9,926 | 4,110 | 17,462 | 820 |
| $t+1$ | 8,426 | 3,326 | 15,032 | 683 |
| $t+2$ | 6,448 | 3,182 | 12,199 | 680 |
| $t+3$ | 4,997 | 2,964 | 10,048 | 679 |
| $t+4$ | 4,036 | 2,618 | 8,269 | 630 |
| $t+5$ | 2,969 | 1,990 | 6,426 | 494 |
| $t+6$ | 833 | 650 | 2,237 | 167 |
| Total | 37,635 | 18,840 | 71,673 | 4,153 |

Table 2.11 – Average real hourly wages (in logs) for treated and control groups

| | N | Mean | St. Dev. |
|---|---|---|---|
| Continuously overqualified | 37,635 | 1.70 | 0.39 |
| Initial overqualified | 18,840 | 1.90 | 0.40 |
| Continuously well-matched | 71,673 | 2.03 | 0.37 |
| Initial well-matched | 4,153 | 1.98 | 0.37 |
| Total | 132,301 | 1.92 | 0.41 |

Table 2.12 – Descriptive statistics of the variables included in the wage regression model (Portugal, 2006-2012)

| Independent Variables | Mean | St. Dev. | Min. | Max |
|---|---|---|---|---|
| **Individual-level characteristics** | | | | |
| Overqualified (OQ) | 0.427 | 0.495 | 0 | 1 |
| Years since entry (YSE) | 3.067 | 1.752 | 1 | 7 |
| Years since entry squared (YSE$^2$) | 12.476 | 12.578 | 1 | 49 |
| Female | 0.640 | 0.480 | 0 | 1 |
| Age25-29 | 0.862 | 0.344 | 0 | 1 |
| Age<25 (omitted category) | 0.138 | 0.344 | 0 | 1 |
| Foreigner | 0.011 | 0.109 | 0 | 1 |
| Education level | | | | |
| Bachelor (omitted category) | 0.09 | 0.285 | 0 | 1 |
| Graduate | 0.877 | 0.329 | 0 | 1 |
| Master | 0.033 | 0.178 | 0 | 1 |
| Field of study | | | | |
| Engineering & technology (omitted category) | 0.149 | 0.356 | 0 | 1 |
| Business | 0.159 | 0.366 | 0 | 1 |
| History, philosophy, etc. | 0.032 | 0.176 | 0 | 1 |
| Arts | 0.018 | 0.135 | 0 | 1 |
| Social sciences | 0.086 | 0.281 | 0 | 1 |
| Education | 0.065 | 0.246 | 0 | 1 |
| Journalism, media studies and communication | 0.022 | 0.146 | 0 | 1 |
| Medicine | 0.147 | 0.354 | 0 | 1 |
| Non-Defined | 0.322 | 0.348 | 0 | 1 |
| Cohort 2007 | 0.506 | 0.499 | 0 | 1 |
| **Job Characteristics** | | | | |
| Type of contract | | | | |
| Permanent contract (omitted category) | 0.543 | 0.498 | 0 | 1 |
| Fixed term contract | 0.444 | 0.496 | 0 | 1 |
| Non-Defined | 0.013 | 0.116 | 0 | 1 |
| Occupation categories 2-digit (ISCO-08): | | | | |
| Legislative power (11, 12, 13, 14) | 0.023 | 0.152 | 0 | 1 |
| Physical science (21, 22, 23, 24, 25, 26) | 0.565 | 0.495 | 0 | 1 |
| Science & engineering (omitted category - 31, 32, 33, 34, 35) | 0.171 | 0.371 | 0 | 1 |
| Administrative staff (41, 42, 43, 44) | 0.175 | 0.380 | 0 | 1 |
| Service and sales (51, 52, 53, 54) | 0.052 | 0.223 | 0 | 1 |
| Skilled construction, industry sector (71, 72, 73, 74, 75) | 0.006 | 0.081 | 0 | 1 |
| Plant, machine operators, assemblers (81, 82, 83) | 0.002 | 0.053 | 0 | 1 |
| Unskilled workers (91, 92, 93, 94) | 0.006 | 0.079 | 0 | 1 |
| Tenure (in months) | 22.751 | 19.332 | 0 | 94 |

Table 2.13 – Wage regressions (Portugal, 2006-2012)

| Dependent variable: log real hourly wages | OLS | Worker FE | Worker & Firm FE |
|---|---|---|---|
| Overqualified ($OQ$) | −0.1530*** | | |
| | (0.0085) | | |
| Years since entry ($YSE$) | 0.1055*** | 0.1281*** | 0.1184*** |
| | (0.0029) | (0.0025) | (0.0027) |
| Years since entry ($YSE$) squared | −0.0097*** | −0.0105*** | −0.0100*** |
| | (0.0004) | (0.0002) | (0.0002) |
| $OQ * YSE$ | 0.0120** | −0.0002*** | 0.0001 |
| | (0.0041) | (0.0004) | (0.0030) |
| $OQ * YSE^2$ | −0.0001 | −0.0002 | 0.0011** |
| | (0.0006) | (0.0004) | (0.0004) |
| Female | −0.0931*** | | |
| | (0.0040) | | |
| Age25-29 (omitted category: age<25) | 0.0238*** | −0.0047 | −0.00004 |
| | (0.0040) | (0.0029) | (0.0026) |
| Foreigner | 0.0656*** | −0.0091 | 0.0008 |
| | (0.0179) | (0.0125) | (0.0019) |
| Education level: (omitted category - bachelor): | | | |
| Graduate | 0.1127*** | 0.0173** | 0.0107 |
| | (0.0058) | (0.0060) | (0.0058) |
| Master | 0.1334*** | 0.0530*** | 0.0383*** |
| | (0.0108) | (0.0097) | (0.0091) |
| Field of study: | | | |
| (omitted category - engineering & technology): | | | |
| History, Philisophy, etc. | 0.0138 | 0.0161 | 0.0222 |
| | (0.0107) | (0.0148) | (0.0156) |
| Arts | 0.0021 | −0.0005 | −0.0098 |
| | (0.0147) | (0.0248) | (0.0263) |
| Business Science | 0.0625 | −0.0039 | 0.0003 |
| | (0.0056) | (0.0082) | (0.0077) |
| Education | −0.0473*** | 0.0508*** | 0.0687*** |
| | (0.0079) | (0.0136) | (0.0162) |
| Journalism, media studies and communication | −0.0987*** | 0.0060 | 0.0428 |
| | (0.0127) | (0.0175) | (0.0207) |
| Medicine | 0.0240*** | −0.0111 | −0.0007 |
| | (0.0050) | (0.0109) | (0.0110) |
| Social Sciences | 0.0463*** | −0.0153 | −0.0139 |
| | (0.0065) | (0.0090) | (0.0085) |
| Cohort 2007 | −0.0322*** | | |
| | (0.0035) | | |
| Fixed term contract | −0.0230*** | −0.0108*** | −0.0198*** |
| | (0.0032) | (0.0023) | (0.0022) |
| $R^2$overall / $R^2$within | 0.239 | 0.826/0.211 | 0.892/0.193 |
| $N$ | 132,301 | 130,886 | 123,853 |

Notes: (i) all regressions include a set of 2-dgit occupation dummies, tenure, non-defined field of study and non-defined contract; (ii) worker-cluster robust standard errors in parentheses; (iii) *, **, and *** denote significant at 10%, 5%, and 1% respectively

Table 2.14 – Wage regressions (Portugal 2006-2012): alternative specification

| Dependent variable: log real hourly wages | OLS | Worker FE | Worker & Firm FE |
|---|---|---|---|
| Overqualified ($OQ$) | $-0.2021^{***}$ | | |
| | (0.0090) | | |
| Years since entry ($YSE$) | $0.1000^{***}$ | $0.1253^{***}$ | $0.1170^{***}$ |
| | (0.0029) | (0.0025) | (0.0027) |
| Years since entry ($YSE$) squared | $-0.0094^{***}$ | $-0.0104^{***}$ | $-0.0100^{***}$ |
| | (0.0003) | (0.0003) | (0.0003) |
| Continuously $OQ * YSE$ | 0.0052 | $-0.0025$ | $-0.0092^{**}$ |
| | (0.0043) | (0.0034) | (0.0033) |
| Continuously $OQ * YSE^2$ | 0.0006 | $0.0010^{**}$ | $0.0018^{***}$ |
| | (0.0006) | (0.0004) | (0.0004) |
| Initial $OQ * YSE$ | $0.0511^{***}$ | $0.0571^{***}$ | $0.0331^{***}$ |
| | (0.0051) | (0.0052) | (0.0048) |
| Initial $OQ * YSE^2$ | $-0.0042^{**}$ | $-0.0044^{**}$ | $-0.0018^{**}$ |
| | (0.0007) | (0.0006) | (0.0006) |
| $R^2$ overall / $R^2$within | 0.2418 | 0.827/0.214 | 0.892/0.195 |
| $N$ | $123,301$ | $130,886$ | $123,853$ |

Notes: (i) the regressions include controls for individual and job characteristics as described in Table 13.
(ii) worker-cluster robust standard errors in parentheses;
(iii) *, **, and *** denote significant at 10%, 5%, and 1%, respectively.

Figure 2.3 – Evolution of hourly wages by match status at entry and subsequent transitions

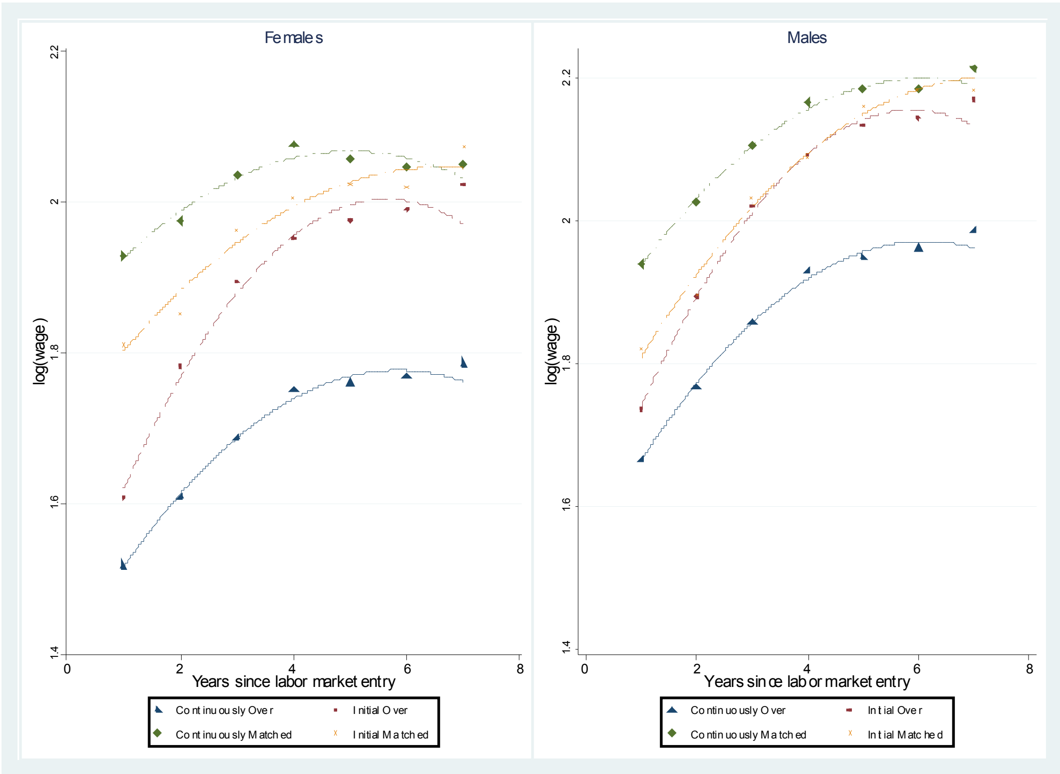

Figure 2.4 – Estimated worker fixed effects

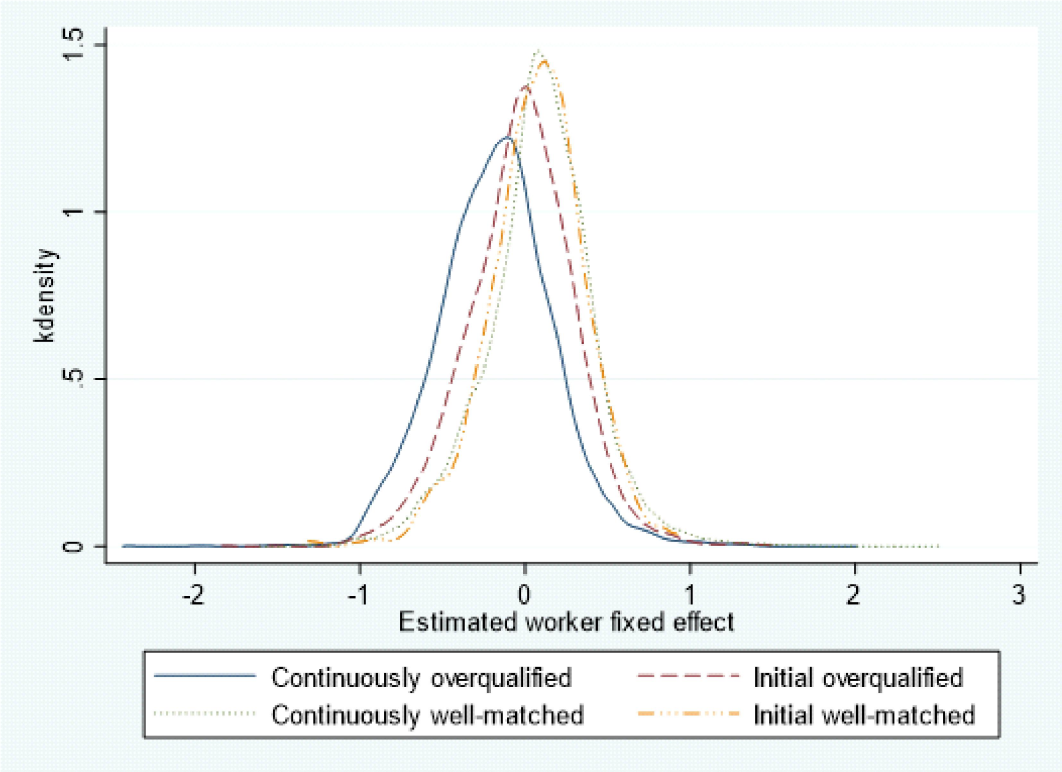

Table 2.15 – APPENDIX A.1. ISCO-08 2-digit Occupation Classification

| ISCO-08 (2-digit) | Occupation categories |
|---|---|
| 11 | "Legislative power and executie bodies representatives, Seniors officials of Public Administration, of special-interest org., entreprises directors, and managers" |
| 12 | "Administration and commercial directors" |
| 13 | "Production and specialised services directors" |
| 14 | "Hotels, food service, trade and others services directors" |
| 21 | "Physical sciences, mathematics, engineering and related techniques specialists" |
| 22 | "Health professionals" |
| 23 | "Teachers" |
| 24 | "Finance, accounting, administrative org., public and trade relations specialists" |
| 25 | "Information and communications technology specialists" |
| 26 | "Legal, social, artistic and cultural ,matters specialists" |
| 31 | "Science and engineering associate professionals" |
| 32 | "Health technicians ans associate professional" |
| 33 | "Financial, business and administration associated professionals" |
| 34 | "Legal, social, sport, cultural and related services, intermediate level technicians" |
| 35 | "Information and communications technicians" |
| 41 | "Office clerks, general secretaries and data keyboard clerks" |
| 42 | "Customer direct support staff" |
| 43 | "Data, accounting, statistical, financial services and material recording operators" |
| 44 | "Other clerical support workers" |
| 51 | "Personal service workers" |
| 52 | "Salespersons" |
| 53 | "Personal care and similar workers" |
| 54 | "Protective and safety services workers" |
| 61 | "Market-oriented farmers and skilled agricultural and farming of animals workers" |
| 62 | "Market-oriented skilled forestry, fishery and hunting workers" |
| 63 | "Subsistence farmers, fishers, hunters and gatherers" |
| 71 | "Building and related trades skilled workers, excluding electricians" |
| 72 | "Metal, machinery and related trades skilled workers" |
| 73 | "Printing and precision instruments manufacturing skilled workers jewelers, craftsman and similar workers" |
| 74 | "Electrical and electronic trades skilled workers" |
| 75 | "Food processing, wood working, garment and other craft, related trade workers" |
| 81 | "Stationary plant and machine operators" |
| 82 | "Assemblers" |
| 83 | "Drivers and mobile plant operators" |
| 91 | "Cleaners and helpers" |
| 92 | "Agricultural, farming of animals, forestry and fishery not skilled workers" |
| 93 | "Minning, construction, manufacturing and transport not skilled workers" |
| 94 | "Food preparation assistants" |
| 95 | "Street vendors (excluding food), and street service workers" |
| 96 | "Refuse workers and other elementary workers" |

Table 2.16 – APPENDIX A.2. Classification of graduate and non graduate jobs by occupation at 2-digit (ISCO-08)

| ISCO-08 at 2-digit | NR-C.A | NR-C.I | RC | RM | NR-M.P | Job Classification | Mode of education (in years) |
|---|---|---|---|---|---|---|---|
| 11 | 1.6112 | 2.3021 | -0.8744 | -1.1852 | -1.3113 | Well-matched | 16 |
| 12 | 1.4161 | 1.6790 | -0.4775 | -1.2805 | -1.3568 | Well-matched | 16 |
| 13 | 1.1725 | 1.9956 | -0.5283 | -0.7799 | -0.8576 | Well-matched | 12 |
| 14 | 0.8870 | 2.1206 | -0.0007 | -0.6433 | -0.7125 | Well-matched | 12 |
| 21 | 1.8330 | 0.5355 | -0.2968 | -0.9816 | -0.9067 | Well-matched | 16 |
| 22 | 1.2768 | 1.0703 | -0.1268 | -0.8338 | -0.5323 | Well-matched | 16 |
| 23 | 1.1878 | 1.6510 | -1.1841 | -1.3690 | -1.2879 | Well-matched | 16 |
| 24 | 1.4964 | 0.8839 | -0.4062 | -1.4755 | -1.5128 | Well-matched | 16 |
| 25 | 1.5313 | -0.3552 | 0.8309 | -0.7611 | -0.8116 | Well-matched | 16 |
| 26 | 1.4444 | 0.3554 | -0.3664 | -1.0821 | -1.1222 | Well-matched | 16 |
| 31 | 0.4181 | -0.0219 | 0.7685 | 0.6341 | 0.5347 | Overqualified | 12 |
| 32 | 0.6943 | 0.4512 | 0.7116 | -0.0649 | -0.0281 | Overqualified | 12 |
| 33 | 0.5005 | 0.1811 | 0.6003 | -0.7882 | -0.8097 | Overqualified | 12 |
| 34 | 0.1737 | 0.3917 | -0.2179 | -0.4523 | -0.2420 | Overqualified | 12 |
| 35 | 0.7429 | -0.7542 | 0.5334 | -0.1180 | -0.2961 | Overqualified | 12 |
| 41 | -0.4103 | -1.1258 | 1.1156 | -0.5818 | -1.0274 | Overqualified | 12 |
| 42 | -0.3134 | -0.2444 | 1.5457 | -0.4210 | -0.9211 | Overqualified | 12 |
| 43 | 0.0011 | -0.5628 | 1.5947 | -0.2388 | -0.7711 | Overqualified | 12 |
| 44 | -0.4875 | -0.9898 | 1.3798 | -0.3124 | -0.6674 | Overqualified | 12 |
| 51 | -0.6836 | 0.2696 | -0.0452 | -0.1230 | -0.1286 | Overqualified | 9 |
| 52 | -0.8447 | -0.0465 | -0.7103 | -0.5301 | -0.6213 | Overqualified | 9 |
| 53 | 0.0656 | 0.5126 | -0.2518 | -0.7084 | -0.5967 | Overqualified | 9 |
| 54 | 0.4098 | 1.0405 | 0.8002 | -0.2754 | 0.6827 | Well-matched | 9 |
| 61 | -0.9095 | -0.7624 | -1.0813 | 0.5696 | 1.1153 | Overqualified | 4 |
| 62 | -0.7173 | -0.3877 | -1.1520 | 0.6063 | 1.2187 | Overqualified | 4 |
| 71 | -0.2670 | -0.0693 | -0.0957 | 0.7846 | 1.3045 | Overqualified | 4 |
| 72 | -0.4406 | -0.6071 | 0.4094 | 1.3140 | 1.2720 | Overqualified | 6 |
| 73 | -0.4484 | -1.2835 | 0.6363 | 1.0716 | 0.5413 | Overqualified | 6 |
| 74 | 0.2938 | 0.1091 | 0.1995 | 0.3035 | 1.2191 | Overqualified | 9 |
| 75 | -0.9220 | -1.0798 | 0.5329 | 1.2715 | 0.7187 | Overqualified | 4 |
| 81 | -0.7998 | -0.7690 | 0.7995 | 2.1755 | 0.9126 | Overqualified | 6 |
| 82 | -0.5707 | -0.7698 | 0.4407 | 1.7236 | 1.121 | Overqualified | 9 |
| 83 | -0.5025 | -0.2652 | 0.3681 | 1.1175 | 1.4862 | Overqualified | 4 |
| 91 | -1.5522 | -1.5267 | -0.4165 | 0.8440 | 0.7375 | Overqualified | 4 |
| 92 | -1.0138 | -0.7880 | -0.9747 | 0.9674 | 1.1311 | Overqualified | 4 |
| 93 | -0.6421 | -0.7176 | 0.3231 | 1.2764 | 1.3710 | Overqualified | 4 |
| 94 | -1.4337 | -0.7091 | 0.2734 | 0.7490 | 0.2171 | Overqualified | 4 |
| 95 | -1.5927 | 0.6046 | -4.0765 | -2.1243 | -1.5618 | Overqualified | 9 |
| 96 | -1.1992 | -1.1628 | 0.6294 | 0.8938 | 1.3202 | Overqualified | 9 |

Table 2.17 – APPENDIX A.3. Wage regressions (Portugal, 2006-2012): excluding initial well-matched workers

| Dependent variable: log real hourly wages | OLS | Worker FE | Worker & Firm FE |
|---|---|---|---|
| Overqualified $(OQ)$ | $-0.1444^{***}$ | | |
| | (0.0091) | | |
| Years since entry $(YSE)$ | $0.0985^{***}$ | $0.1126^{***}$ | $0.1155^{***}$ |
| | (0.0030) | (0.0025) | (0.0028) |
| Years since entry $(YSE)$ squared | $-0.0088^{***}$ | $-0.0105^{***}$ | $-0.0100^{***}$ |
| | (0.0004) | (0.0.0003) | (0.0002) |
| $OQ * YSE$ | $0.0181^{***}$ | $0.0129^{***}$ | 0.0007 |
| | (0.0041) | (0.0032) | (0.0030) |
| $OQ * YSE^2$ | $-0.0010$ | $-0.0002$ | $0.0011^{**}$ |
| | (0.0006) | (0.0004) | (0.0004) |
| $R^2$overall / $R^2$within | 0.252 | 0.830/0.208 | 0.893/0.192 |
| $N$ | 128, 148 | 126, 733 | 119, 970 |

Notes: (i) all regressions include controls for individual and job characteristics as defined in Table 13;
(ii) worker-cluster robust standard errors in parentheses;
(iii) *, **, and *** denote significant at 10%, 5%, and 1%, respectively

Table 2.18 – APPENDIX A.4. Wage regressions (Portugal, 2006-2012): excluding initial overqualified workers and initial well-matched workers

| Dependent variable: log real hourly wages | OLS | Worker FE | Worker & Firm FE |
|---|---|---|---|
| Overqualified ($OQ$) | $-0.1961^{***}$ | | |
| | (0.0117) | | |
| Years since entry ($YSE$) | $0.0974^{***}$ | $0.1197^{***}$ | $0.1089^{***}$ |
| | (0.0031) | (0.0026) | (0.0029) |
| Years since entry ($YSE$) squared | $-0.0088^{***}$ | $-0.0104^{***}$ | $-0.0101^{***}$ |
| | (0.0004) | (0.0002) | (0.0003) |
| $OQ * YSE$ | $0.0184^{***}$ | 0.0005 | $-0.0078$ |
| | (0.0045) | (0.0034) | (0.0033) |
| $OQ * YSE^2$ | $-0.0011$ | 0.0010 | $0.0019^{***}$ |
| | (0.0006) | (0.0005) | (0.0004) |
| $R^2$ overall / $R^2$ within | 0.2574 | 0.841/0.170 | 0.896/0.171 |
| $N$ | 109,308 | 107,893 | 102,402 |

Notes: (i) all regressions include controls for individual and job characteristics as defined in Table 13.
(ii) worker-cluster robust standard errors in parentheses;
(iii) *, **, and *** denote significant at 10%, 5%, and 1%, respectively.

# Chapter 3

# Overqualification in Early Career and Future Job Mobility

Abstract: Job change decisions in early career may have a great impact in the worker's future career prospect. In this paper, we propose to compare the mobility pattern of recent graduates who entered the labor market for the first time in 2006 or 2007 overqualified with the mobility pattern of recent graduates who entered the labor market for the first time in 2006 or 2007 well-matched. The data indicate that the early career of young graduate workers in Portugal is characterized by low job mobility. In fact, we found that more than half of the graduates who entered the labor market for the first time in 2006 or 2007 did not change their first employer over the analysed period. Using the number of employer changes as an indicator of job mobility, the econometric approach revealed that there is no statistical difference in terms of the total number of job changes experienced by overqualified workers at entry and well-matched workers at entry. However, when we consider a finer measure of job mobility - the total number of job changes to a well-matched job - the estimates indicate that overqualified workers at entry are less likely to switch to a well-matched job. In other words, a recent graduate who entered the first job overqualified experienced, on average, 35% less number of job changes to a well-matched job than a similar graduate who entered the first job well-matched.

KEYWORDS: Job mobility, skill mismatches, overqualification, wages.

JEL CODES: J31, I23, J62, J63

## 3.1 Introduction

Younger workers, new entrants in the labor market, are generally more prone to job changes as in early career younger workers search for a good job match (e.g., Jovanovic, 1979). Decision to quit a job is not taken lightly. In fact, job change decisions at the start of the career are crucial in terms of future career prospects and their consequences are long lasting (e.g., Argaw *et al.,* 2017). In this process of searching for a good match, models of job change offer an explanation for the reasons why younger workers are more job mobile during the beginning of their career.

According to the job search model, the wage in the current job may be a determinant of worker turnover. As such, workers decisions to change job are based on the comparison between their wages and the ones offered in the new job (e.g., Mortensen, 1986). Several extensions of the original model were made, for example, Burdett (1978) extended the classical job search model, the so-called model of worker search, by allowing workers to search for a better job while employed.

The matching theory also offers a contribution to explain job mobility. According to the matching theory, overqualification is a temporary mismatch and is associated with imperfect information and job search costs. As time goes by, workers' job tenure increases and job search diminishes. Johnson (1978) provides a theoretical model of mobility and find evidence that education is negatively related to job mobility and to job tenure. According to this author, job mobility rates are much higher for workers who lack experience. In this line of thought, Jovanovic (1979) predicts a model where workers mobility depends on the degree of their current jobs' productivity: workers change job if the productivity of their current job is low. The main conclusion is that job mobility will depend on whether there is a good or a bad match between the worker and the firm (Sicherman, 1991).

The career mobility theory (Sicherman, 1991) also reveals that overqualification is a temporary phenomenon. According to this line of thought, overqualification serves to offset the lack of experience and training in the early career. As those gaps in training are filled, workers are able to change to well-matched jobs. However, if overqualification is associated with individual's unobserved characteristics, the mismatch may be more persistent or even permanent.

The decision to search for a better match can be based on monetary reasons or based on nonwage factors (e.g., commuting time, working conditions, job security). Authors such as Hwang *et al.* (1998), Bonhomme and Jolivet (2009), Sullivan and To

(2014) or Bonhomme *et al.* (2016) contribute to this literature and highlighted the importance of nonwage characteristics in workers' decision to change job. They argue that wages are not enough to explain workers' decision to accept a job offer, inasmuch as other nonwage factors are important at the moment of decision making. Nonwage job characteristics such as, commuting time (White, 1992), promotions (e.g. Bonhomme *et al.*, 2016) or flexible hours (Altonji and Paxson, 1992) may also affect turnover rates. More recently, Argaw *et al.* (2017) extended the model proposed by Sullivan and To (2014) and incorporated risk attitudes to nonwage job characteristics when workers decide whether or not to accept a job offer.

It is at the start of a career, where investments in specific capital are smaller, that workers invest in this process of finding the best match. For overqualified workers this search will be more relevant as, everything else remaining constant, the worker is, under this point of view, in a job that is maladjusted to his qualifications. Overqualification may arise from several imbalances between the skills offered and demanded in the labor market affecting productivity, turnover rates (Sicherman, 1991; Hersch, 1991) and job satisfaction (e.g., Maynard, Joseph, and Maynard, 2006). Hence, overqualified workers may have surplus skills, abilities, experience, qualifications than the one they are supposed to have to perform their jobs (Erdogan, Bauer, Peiró, and Truxillo, 2011).

Regarding job mobility, overqualified workers seem to be more likely to change job. The first reason may have to do with wage penalty. In fact, studies that have focused on the earnings consequences of overeducation (Duncan and Hoffman, 1981; Hartog and Oosterbeek, 1988; Verdugo and Verdugo, 1989; Groot, 1993; Cohn and Kahn, 1995; Kiker, Santos and Oliveira, 1997; Battu *et. al*, 1999) found that when compared with their job co-workers who are adequately educated, overeducated workers receive a wage bonus for the extra years of surplus schooling, even though smaller than the returns to required education and that they earn less than their counterparts with the same years of schooling, but who are well-matched. Second, overqualified workers may be more likely to change job because they are working in an inadequate job, they may be frustated with the lack of career opportunities or by the fact that they are working in a field unrelated to their education (Maynard, Joseph, and Maynard, 2006). Thus, overeducation seems to be associated with lower job satisfaction and, consequently, to lower job productivity (e.g., Tsang *et al.*, 1991). In their study, Battu *et al.* (1999) use a survey of graduates from two cohorts, 1985 and 1990, and examine the determinants of overeducation in the UK. They argue that individual's job satisfaction and earnings are adversely affected by overqualification over time, as such, well-matched workers report higher levels of job satisfaction. Fleming *et al.* (2008), using data from the first wave of the Household, Income and Labour Dynamcis in Australia, find that overeducated workers are less satisfied than their counterparts who are well-matched. In general, workers that are more mobile are the ones that receive lower wages and exhibit lower job satisfaction (McGuinness and Sloane, 2011).

The literature on the relationship between job mobility and overqualification is still scarce. Even though, there are few studies that in particular analyse the relationship between job mobility and overqualification status at entry. For example, Feldman and Turnley (1995) discuss the turnover rates of overqualified workers and find that they report higher intentions to quit their jobs when compared to non-underemployed workers, one reason is that, being overqualified in the early career has the same negative psychological consequences as unemployment for a recent graduate. In a recent study, Meroni *et al.* (2017) using the 2005 REFLEX data (REsearch into employment and professional FLEXibility) find evidence that overeducation in early career leads to higher chances of being overeducated later on. In the same spirit, Voßemer *et al.* (2016) using data from the German SOEP for the period 1984–2012, find evidence that unemployed workers who re-enter the labor market overeducated delay their transition to a well-matched job. Argaw *et al.* (2017) using SOEP data find that risk-averse individuals are the ones who change their jobs less often during early career. Topel and Ward (1992), using data based on the Longitudinal Employee-Employer Data (LEED) from 1957 to 1972, find that during the first ten years in labor market young workers hold, on average, seven jobs, five of them during the first five years. According to them, this period is characterized by a transition to relatively stable employment and earnings growth. von Wachter and Bender (2006), using data for Germany, focus on the longterm effects of the job loss in young worker's early career and find that their wage losses are initially 15 percent, but drop to zero within five years.

To the best of our knowledge, there are few studies that provide evidence on the impact of overqualification on younger workers' future mobility pattern. As such, the aim of this paper is to extend the literature on job mobility by comparing the mobility pattern of the recent graduates who entered for the first time in the labor market overqualified with graduates who entered the labor market for the first time in a well-matched job. Using a sample of recent graduates from two cohort years (2006 and 2007) and a Poisson regression, this paper contributes to the literature on skill mismatches by examining to what extent overqualified workers change jobs more often during early career in searching for an adequate job than well-matched workers. The rest of the paper is organized as follows. In Section 3.2 we describe the data and sample construction. Empirical results are presented and discussed in Section 3.3. Section 3.4 concludes.

## 3.2 Data and Methodological Issues

### 3.2.1 Quadros de Pessoal (QP) and O*NET

Our data come from *Quadros de Pessoal (QP),* a large longitudinal matched employer-employee dataset from the Portuguese Ministry of Labor, Solidarity, and Social Security. All firms in the private sector employing at least one wage earner are legally obliged to fill in this survey.[1] The data include yearly information at the firm/establishment level (e.g., industry, location, employment, economic activity), and worker level (e.g., gender, age, education, occupation, qualification, tenure, wages, hours worked). All firms, establishments, and workers entering *QP* dataset have a unique identifying number, so they can be followed across all annual waves of data. Furthermore, the worker files include the firm and establishment number to which each individual is affiliated in a given year, allowing to match workers with their employers both at the firm and the establishment level.

We also use data from the Occupational Information Network database (O*NET, version 21.0), database that contains a rich set of variables that describe job requirements and worker competencies required to perform those jobs.[2] O*NET is a successor of the U.S. Department of Labor's Dictionary of Occupational Titles (DOT), a database that imputes to workers the tasks measures associated with their occupations. According to Altonji *et al.* (2014), the O*NET database is very useful to measure tasks importance within occupations. However, this analysis is not without limitations. The first limitation is that we assume that workers are effectively using these tasks in their jobs, the second is that occupational task measures are fixed over time. A third limitation is that both, DOT and O*NET contain numerous potential task scales, however it turns very difficult to understand which measure best represents a given task construct (Acemoglu and Autor, 2011). We follow Acemoglu and Autor (2011) and use composite

[1] Public administration, self-employment, and nonmarket services are not covered by QP.

[2] We are very thankful to Miguel Portela for sharing data and stata codes.
For further information, please see https://www.onetonline.org.

measures of O*NET Work Activities and Work Context Importance scales and group tasks into five categories, according to the intensity of their use in a given occupation (see Table 3.1 for details).

*INSERT TABLE 3.1 HERE*

The five categories are: (i) non-routine cognitive: analytical task (NR-C.A); (ii) non-routine cognitive: interpersonal task (NR-C.I); (iii) routine cognitive task (RC); (iv) routine manual task (RM); (v) non-routine manual physical task (NR-M.P). These data were collected using the O*NET-SOC occupational Classification scheme which we collapse into SOC occupations, following the methodology used by Acemoglu and Autor (2011).[3] Finally, we recode and normalized the O*NET data into a two-digit ISCO-08 coding by applying a crosswalks between them (data and codes prepared by the Institute for Structural Research, www.ibs.org.pl/resources). Each scale is then standardized to have mean zero and standard deviation one, i.e., we compute the relative importance of tasks within a two-digit occupation code (Acemoglu and Autor, 2011).

We restrict our analysis to the period between 2006 and 2012 and to newly graduates who entered the labor market for the first time in 2006 and 2007.[4] As such, we select workers with a tenure equal to zero in the first job and who were absent from *QP* files in the last five or six years before their entry in the labor market. We further exclude from our analysis workers in public administration, agricultural and fishery (occupations 11, 61 and 92) and workers aged above 30 years old at the time of their entry in the labor market. Our sample is composed of 32,305 workers who enter for the first time in the labor market in 2006 or 2007. Moreover, we use some correction routines to avoid potential biases such as the repair of missing and inconsistent values in gender, age, and education, otherwise we dropped them.[5] In 2010, the Portuguese Classification of Occupations (*CPP*) changed in the sense of grouping together specific occupations. Whenever possible we recode occupations according to their new classification, at the

[3] The Standard Occupational Classification contains links to major groups, the complete hierarchical structure, broad and detailed occupational definitions (https://www.bls.gov).

[4] Newly graduates include Bachelors, Graduates, and Master's degree. We exclude doctorates from our sample due to their lack of representativeness (Santos *et al.*, 2016; Almeida *et al.*, 2017).

[5] The top and bottom 1 percent observations in wages were excluded from the sample. Multiple job-holder workers were also dropped.

three-digit level so we can match them with ISCO-08 classification. We also recode the Industry Classification under the ISIC Rev. 3 whenever possible.

### 3.2.2 Measuring Overqualification and Job Mobility

The first variable of interest used in the present study defines whether a worker is overqualified in his first job or not, according to the intensity of the use of the five categories of tasks in the current occupation. First, we define a graduate job as an occupation where recent graduate workers are well-matched. In other words, a graduate job refers to a 2-digit occupation where the value of the standardized task "non-routine cognitive analytical" or "non-routine cognitive interpersonal" is greater than one. Whenever both values are smaller than one, the worker is overqualified in his occupation. In Appendix A.5. we present the classification of occupations categories at a two-digit level (ISCO-08), and in Appendix A.6. we present the value of the standardized tasks within each two-digit occupation code (ISCO-08). The latter allows us to classify occupations as graduate or non graduate job, in other words, occupations where newly graduate workers are well-matched or overqualified. According to our classification, workers in occupations that correspond to major groups 1 and 2 as well as to occupation code 54 are considered well-matched. Overqualified graduates are in occupations that correspond to the remaining major groups.

*INSERT APPENDIXES A.5. AND A.6. HERE*

For example, working in occupation with ISCO-08 code 41 "Office clerks, general secretaries and data keyboard clerks" requires a low intensity of non routine cognitive analytic tasks (-0.41), non routine cognitive interpersonnal tasks (-1.13), routine manual tasks (-0.58) and non routine manual physical tasks (-1.03) and requires a high intensity of routine cognitive tasks (1.12). As such, a newly graduate that enters for the first time in the labor market in this occupation is considered to be overqualified inasmuch as the value of "non-routine cognitive: analytical" and "non-routine cognitive: interpersonal" is lower than one. Using the mode of education at a 2-digit level, graduate jobs (the mode of education is at least 15 years of schooling) correspond to major groups 1 and

2, with the exception of occupations 13 and 14.[6] Overeducated workers are working in the remaining major groups (occupations that do not require a graduation and whose mode of education is lower than 15 years).

The second variable of interest describes job mobility and corresponds to the total number of job changes that a recent graduate worker experienced during the analysed period. In our approach, a job change corresponds to a change in employer. Thus, a job change occurs when the firm identifying number where the worker is employed in year $t$ is different from the identifying number where the worker is employed in year $t-1$ or in the last year the worker appears in the *QP* files. Unfortunately, we cannot say for sure if the change is voluntary or not, as we are not able to provide this distinction in *QP*.

Our paper focuses heavily on recent graduates who entered the labor market for the first time. As such we need to define "early career" in terms of the number of years following labor market entry. Topel and Ward (1992) consider that "early career" corresponds to the first ten years after entry. Neumark (2002) defines "early career" as "the five-year post-schooling period", whereas Argaw *et al.* (2017), more recently, define "early career" as the first seven years following labor market entry. In our study we define "early career" as the first five or six years following labor market entry, according to the respective cohort of workers, 2006 or 2007.

### 3.2.3 Descriptive Statistics

Table 3.2 presents summary statistics of workers who enter for the first time in the labor market overqualified or well-matched (by gender). 14,024 of a total of 32,305 workers are overqualified (about 43.4%) and 18,281 out of this total are well-matched (about 56.6%) in their first job. Furthermore, in our data, newly graduate workers are predominantely women. Almost 63% of the total number of workers (overqualified and well-matched) are women.

*INSERT TABLE 3.2 HERE*

[6] This result is consistent with previous studies that found that a larger proportion of entrepreneurs/directors in Portugal has lower levels of education (e.g., Rocha *et al.*, 2015).

In Table 3.3 we report some comparative statistics for recent graduates workers according to their levels of schoolling. The great majority of both overqualified and well-matched workers have a graduation. Almost 83% (78%) of overqualified female (male) workers have a graduation whereas almost 90% (85%) of well-matched female (male) workers have the same degree. Moreover, there is a larger percentage of bachelors in the group of overqualified workers when compared with the group of well-matched workers.

*INSERT TABLE 3.3 HERE*

Table 3.4 reports the total number of job changes (NJC), by gender and match status at entry. The data show that more than half of the newly graduates who enter the labor market in 2006 or 2007 did not change their first employer over the analysed period. In fact, approximately 55 percent (59 percent) of females (males) who were overqualified at entry and 60 percent (58 percent) of females (males) who were well-matched at entry did not change their first employer over the analysed period. Furthermore, only one third of recent graduates who were overqualified at entry changed employer at least once over the analysed period.

*INSERT TABLE 3.4 HERE*

Table 3.5 shows the total number of job changes to a well-matched job (NJCWM), by gender and match status at entry. The data show that the great majority of females (males) who entered the first job overqualified, 85 percent (80 percent), did not move to a well-matched employer throughout the analysed period. The table also shows that only 13 percent (17 percent) of overqualified females (males) changed once to a well-matched job. This proportion nearly doubles for workers who were well-matched at entry, in fact 26 percent of both females and males have changed employer once to a well-matched job during the analysed period. Moreover, a very few proportions of workers have changed twice to a well-matched job, for example, only 2 percent of overqualified female workers have changed twice to a well-matched employer.

*INSERT TABLE 3.5 HERE*

These results seem to confirm that the Portuguese labor market is characterized by low job mobility of younger workers, especially at a time of economic crisis, and differ from the results obtained in the literature. Topel and Ward (1992) find that workers, on average, hold seven jobs within ten years in the labor market whereas von Watcher and Bender (2006) argue that the early career of young workers are characterized by high job mobility.

## 3.3 Job Mobility

### 3.3.1 Empirical Strategy

The aim of this essay is to analyse to what extent the mobility pattern of recent graduates who entered the labor market for the first time in 2006 or 2007 overqualified differ from the mobility pattern of similar graduates who entered the labor market in a well-matched job. To examine this relationship we use, in the empirical analysis, a sample of 32,305 individuals who entered the labor market for the first time in 2006 or 2007. The model writes as:

$$NJC_i = \alpha_0 + \alpha_1 OQ_{entry_i} + \delta \mathbf{X}_{entry_i} + \varepsilon_i \tag{3.1}$$

Where $NJC_i$ is the total number of job changes that occurred during the analysed period for individual $i$. $OQ_{entry}$ is a binary variable which takes the value one if a worker is overqualified in the first job and 0 if a worker is in a well-matched job in the first job. $\mathbf{X}_{entry}$ is a set of control variables for worker, firm and job characteristics at entry, and $\varepsilon_i$ is a random error term assumed to be uncorrelated with the regressors. As our dependent variable is a count variable we use a Poisson regression to estimate equation (1), as OLS can generate inaccurate predictions (e.g., negative counts), be inefficient or even biased. The use of a Poisson regression makes sense when the dependent variable takes low values and possibly with many zeros. Poisson regression assumes that each observed count $y_i$ follows a Poisson distribution with parameter $\lambda_i$. The Poisson model has the property that mean and variance are equal (equidispersion property):[7]

[7] The consequences of overdispersion are similar to those of heteroscedasticity in the linear regression model. If we apply a Poisson regression to overdispersed data the estimates for the standard deviation of the coefficients will be biased towards zero meaning that the z statistics will be inflated. As for many count data the variance exceeds the mean (overdispersion), we used a Negative Binomial Regression that allows the variance to differ from the mean. According to the likelihood ratio test of $\alpha = 0$, the errors do not exhibit overdispersion, thus the NB regression model is rejected in favor of the Poisson regression model. Furthermore, we follow Long and Freese (2014) procedure and also found that the

$$E(y) = \lambda \text{ and } V(y) = \lambda$$

The description of the independent variables included in the model is presented in Table 3.6

*INSERT TABLE 3.6 HERE*

To provide an overview of the sample, the descriptive statistics of the variables included in the model are reported in Table 3.7 Almost 44% of recent graduates who entered the labor market for the first time in 2006 or 2007 are overqualified. The great majority of recent graduates are females (63,4%), have a fixed term contract (67,2%), possess a degree (84,4%) and are located in a urban area (76.7%).

*INSERT TABLE 3.7 HERE*

### 3.3.2 Empirical Results

Table 3.8 provides estimation results for equation (1) using the Poisson model (PRM) and, for comparison purposes, the ordinary least squares model (OLS). Furthermore, in the last two columns we run the same models, considering now as the dependent variable the total number of job changes to a well-matched job (NJCWM) as defined in section 3.2.

*INSERT TABLE 3.8 HERE*

The Poisson estimates of equation (1) are reported in column (1) and indicate that, ceteris paribus, at entry, there is no significant difference between being overqualified at the begin of a career or well-matched in terms of total number of job changes during the analysed period. In this regard, the mobility pattern of overqualified workers is very similar to the mobility pattern of well-matched workers. Males, older workers, those who hold a fixed term contract and whose field of study is business science changed jobs more often during the analysed period than females, younger workers, those with

Poisson model is the model that fits better the data.

a permanent contract and engineers. Furthermore, receiving higher wages at entry decreases the chances to change job. Looking at the wage dispersion, we can observe that the greater the wage dispersion in a given occupation, the greater the probability to change job. Workers employed in larger firms experience, on average, a small number of job changes, while workers located in an urban area experience, on average, a larger number of job changes. The OLS results reported in column (2) are qualitatively identical.

The Poisson estimates of equation (1) considering now as the dependent variable the total number of job changes for a well-matched job are reported in column (3). We can observe that the independent variable "Overqualified (OQ)" is now statistically significant. The Poisson estimate indicates that an overqualified worker at entry is expected to change less to a well-matched job than a similar worker who entered in a first job well-matched. An individual who enters the first job overqualified experience, on average, 35% less number of job changes to a well-matched job when compared with a similar individual who enters the first job well-matched.[8] Being a female decreases the number of changes to a well-matched job by 13%, holding all other variables constant. On contrary, having a degree increases the chance to change to a well-matched job in comparison to a bachelor. Portuguese natives, males and older workers are expected to change more to a well-matched job than other workers. The latter result seems to suggest that young workers need some time to gain experience and to signal their skills to be able to change to a well-matched job. Moreover, those that earn higher wages are more likely to change to a well-matched job. Thus, individuals who enter in a first job with higher wages experience, on average, a small number of job changes but face a larger chance to change to a well-matched job, suggesting that this group of workers may correspond to a higher-ability group of workers. The OLS results in column (4) indicate that, qualitatively, the results are very similar.

In sum, our results reveal that overqualified workers exhibit a similar mobility pattern in terms of total number of job changes. However, when we consider the number of job changes for a well-matched job, being overqualified at entry is predicted to change less to a well-matched job.

[8] The exact value is computed $[(\exp(\hat{\beta}) - 1) * 100]$.

## 3.4 Concluding Remarks

In this paper, we examine the mobility pattern of recent graduates who entered for the first time in the labor market, in 2006 or 2007. Exploring a large matched employer-employee data set over the 2006-2012 period, we constructed two indicators of job mobility based on the number of employer changes to compare the mobility pattern of recent graduates who entered the labor market overqualified with the mobility pattern of recent graduates who entered the labor market well-matched.

Our data seem to suggest that the early career of young graduate workers in Portugal is characterized by low job mobility. In fact, we found that more than half of the newly graduates who entered in the labor market for the first time in 2006 or 2007 did not change their first job. Furthermore, we also found that the vast majority of workers who entered in the labor market overqualified were not able to move to a well-matched job throughout the analysed period.

Using the total number of employer changes over the analysed period as an indicator of job mobility, the econometric approach revealed no statistical differences in terms of the total number of job changes experienced by overqualified workers at entry and well-matched workers at entry. However, when we consider a finer measure of job mobility - the total number of job changes to a well-matched job - the estimates indicate that overqualified workers at entry are less likely to switch to a well-matched job. More precisely, a recent graduate who entered the first job overqualified experienced, on average, 35% less number of job changes to a well-matched job than a similar graduate who entered the first job well-matched.

Furthermore, the regression results revealed that females, foreigners and younger workers, are less likely to change to a well-matched job. Workers that hold a degree are predicted to change to a well-matched job more often than their similar bachelor counterparts. In general, workers in engineering and technology are more likely to switch to a well-matched job when compared with similar workers in other fields of study (the only exception is medicine). Regarding job and firm characteristics, new entrants employed in smaller firms and holding a fixed term contract are more likely to change job and, in particular, to change to a well-matched job. Initial wages seem to be a good predictor of future mobility. In particular, we found that workers with higher wages in first job, experience, on average, few job changes, but when they move they are more likely to switch to a job that matches their skills.

# Bibliography

[1] Acemoglu, D., Autor, D. (2011). "Skills, tasks and technologies: Implications for employment and earnings". Elsevier, Part B of Handbook of Labor Economics, 4: 1043-1171.

[2] Almeida, A., Figueiredo, H., Cerejeira, J., Portela, M., Sá, C., Teixeira, P. (2017). "Returns to Postgraduate Education in Portugal: Holding on to a Higher Ground?". IZA DP No. 10676.

[3] Altonji, J. G., Paxson, C. H. (1992). "Labor supply, hours constraints, and job mobility". Journal of Human Resources, 27(2): 256-278.

[4] Altonji, J. G., Kahn, L. B., Speer, D. J. (2014). "Trends in earnings differentials across College Majors and the changing task composition of jobs." American Economic Review. American Economic Association, 104(5): 387-393.

[5] Argaw, B. A., Maier, M. F., Skriabikova, O. J. (2017). "Risk attitudes, job mobility and subsequent wage growth during the early career". ZEW Discussion Papers, Nº17-023, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

[6] Battu, H., Belfield, C. R., Sloane, P. J. (1999). "Overeducation among graduates: a cohort view". Education Economics, 7(1): 21-38.

[7] Bonhomme, S., Jolivet, G. (2009). "The pervasive absence of compensating differentials". Journal of Applied Econometrics, 24(5): 763-795.

[8] Bonhomme, S., Jolivet, G., Leuven, E. (2016). "School characteristics and teacher turnover: assessing the role of preferences and opportunities". The Economic Journal, 126(594): 1293-1633.

[9] Burdett, K. (1978). "A Theory of employee job search and quit rates". The American Economic Review, 68(1): 212-220.

[10] Cohn, E., Khan, S.P. (1995). "The wage effect of overschooling revisited". Labour Economics, 2(1): 67-76.

[11] Duncan, G. J, Hoffman, S.D. (1981). "The incidence and wage effects of overeducation". Economics of Education Review, 1(1): 75-86.

[12] Erdogan, B., Bauer, T. N., Peiró, J. M., Truxillo, D. M. (2011). "Overqualified employees: Making the best of a potential bad situation for individuals and organizations". Industrial and Organizational Psycology, 4(2): 215-232.

[13] Feldman, D. C., Turnley, W. H. (1995). "Underemployment among recent business college graduates". Journal of Organizational Behavior, 16: 691-706.

[14] Fleming, C. M., Kler, P. (2008). "I'm too clever for this job: a bivariate probit analysis on overeducation and job satisfaction in Australia". Applied Economics, 40(9): 1123-1138.

[15] Groot, W. (1993). "Overeducation and the returns to enterprise-related schooling". Economics of Education Review, 12(4): 299-309.

[16] Hartog, J., Oosterbeek, H. (1988). "Education, allocation and earnings in the Netherlands: Overschooling?". Economics of Education Review, 7(2): 185-194.

[17] Hersch, J. (1991). "Education match and job match". Review of Economics and Statistics, 73(1): 140-144.

[18] Hwang H., Mortensen, D., Reed, W. (1998). "Hedonic wages and labor market search". Journal of Labor Economics 16: 815–847.

[19] Johnson, W. R. (1978). "A Theory of job shopping". The Quarterly Journal of Economics, 92(2): 261-278.

[20] Jovanovic, B. (1979). "Job matching and the theory of turnover". Journal of Political Economy, 87(5), Part 1: 972-990.

[21] Kiker, B.F., Santos, M.C., De Oliveira, M. M. (1997). "Overeducation and undereducation: evidence for Portugal". Economics of Education Review, 16(2): 111-125.

[22] Long, J. S., Freese, J. (2014). "Regression Models for Categorical Dependent Variables Using Stata, Third Edition". Stata Press.

[23] Maynard, D. C., Joseph, T. A., Maynard, A. M. (2006). "Underemployment, job attitudes, and turnover intensions". Journal of Organizational Behavior, 27(4): 509-536.

[24] McGuinness, S., Sloane, O. J. (2011). "Labour Market Mismatch among UK Graduates: An Analysis using REFLEX data". Economics of Education Review, 30(1): 130-145.

[25] Meroni, E. C., Vera-Toscano, E. (2017). "The persistence of overeducation among recent graduates". Labour Economics, 48: 120-143.

[26] Mortensen, D. (1986). "Job Search and Labor Market Analysis" in Handbook of labor economics, Edited by O. Ashenfelter and R. Layard, Volume II: 849-919.

[27] Neumark, D. (2002). "Youth labor markets in the united states: Shopping around vs. staying put". The Review of Economics and Statistics, 84(3): 462-482.

[28] Rocha, V., Carneiro, A., Varum Amorim, C. (2015). "Entry and exit dynamics of nascent business owners". Small Business Economics, 45(1): 63-84.

[29] Santos, J. M., Horta, H., and Heitor, M. (2016). "Too many phds? an invalid argument for countries developing their scienti c and academic systems: The case of portugal". Technological Forecasting and Social Change, 113, Part B: 352-362.

[30] Sicherman, N. (1991). "Overeducation in the Labor Market". Journal of Labor Economics, 9(2): 101-122.

[31] Sullivan, P., To, T. (2014). "Search and Nonwage Job Characteristics". The Journal of Human Resources, 49(2): 472-507.

[32] Topel, R. H., Ward, M. P. (1992). "Job mobility and the careers of young men". The Quarterly Journal of Economics, 107(2): 439-479.

[33] Tsang, M. C., Rumberger, R. W., Levin, H. M. (1991). "The impact of surplus schooling on worker productivity". Industrial Relations, 30(2): 209-228.

[34] Verdugo, R., Verdugo, N. (1989). "The impact of surplus schooling on earnings". Journal of Human Resources, 24(4): 629-643.

[35] Voßemer, J., Schuck, B. (2016). "Better overeducated than unemployed? The short and long term effects of an overeducated labour market re-entry". European Sociological Review, 32(2): 251–265.

[36] von Watcher, T., Bender, S. (2006). "In the right place at the wrong time: The role of firms and luck in young workers' careers". American Economic Review: 96(5): 1679-1705.

[37] White, K. L. (1992). "The determinants of Black-White differences in early employment careers: search, layoffs, quits, and endogenous wage growth". The Journal of Political Economy, 100(3): 535-560.

# Tables and Figures

Table 3.1 – O*NET Data descriptors (Acemoglu and Autor, 2011)

| Classification | Tasks |
|---|---|
| Non-routine Cognitive Analytical (NR-C.A) | Analyzing Data or Information<br>Thinking Creatively<br>Interpreting the Meaning of Information for Others |
| Non-routine Cognitive Interpersonal (NR-C.I) | Coaching and Developing Others<br>Guiding, Directing, and Motivating Subordinates<br>Establishing and Maintaining Interpersonal Relationships |
| Routine Cognitive (RC) | Importance of Being Exact or Accurate<br>Importance of Repeating Same Tasks<br>Structured versus Unstructured Work |
| Routine Manual (RM) | Pace Determined by Speed of Equipment<br>Spend Time Making Repetitive Motions<br>Controlling Machines and Processes |
| Non-routine Manual physical (NR-M.P) | Spatial Orientation<br>Manual Dexterity<br>Operating Vehicles, Mechanized Devices, or Equipment<br>Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls |

Table 3.2 – Distribution of overqualified and well-matched workers by cohort and gender

| | Overqualified | | | Well-matched | | |
|---|---|---|---|---|---|---|
| | Males | Females | Total | Males | Females | Total |
| Cohort 2006 | 2,149 | 4,315 | 6,464 | 3,386 | 5,804 | 9,190 |
| Cohort 2007 | 2,580 | 4,980 | 7,560 | 3,707 | 5,384 | 9,091 |
| Total | 4,729 | 9,295 | 14,024 | 7,093 | 11,188 | 18,281 |

Table 3.3 – Distribution of overqualified and well-matched workers at entry by level of schooling

| | Overqualified (%) | | Well-matched (%) | |
|---|---|---|---|---|
| Education level | Males | Females | Males | Females |
| Bachelors | 20.7 | 15.8 | 13.5 | 8.8 |
| Graduates | 77.6 | 82.6 | 84.7 | 89.7 |
| Masters | 1.7 | 1.6 | 1.8 | 1.5 |
| Total | 100% | 100% | 100% | 100% |

Table 3.4 – Distribution of workers by NJC, by gender and match status at entry

| Distribution of workers by NJC, by gender and match status at entry | | | | | | |
|---|---|---|---|---|---|---|
| | Overqualified | | | Well-matched | | |
| | Males | Females | Total | Males | Females | Total |
| 0 | 2,812 (59.4%) | 5,146 (55.4%) | 7,958 | 4,121 (58.1%) | 6,751 (60.3%) | 10,872 |
| 1 | 1,481 (31.3%) | 3,158 (34%) | 4,639 | 2,246 (31.7%) | 3,453 (30.9%) | 5,699 |
| 2 | 380 (8%) | 840 (9%) | 1,220 | 602 (8.5%) | 827 (7.4%) | 1,429 |
| 3 | 48 (1%) | 130 (1.4%) | 178 | 117 (1.6%) | 143 (1.3%) | 260 |
| 4 or + | 8 (0.3%) | 21 (0.2%) | 29 | 7 (0.1%) | 14 (0.1%) | 21 |
| Total | 4,729 | 9,295 | 14,024 | 7,093 | 11,188 | 18,281 |

Table 3.5 – Distribution of workers by NJCWM, by gender and match status at entry

| Distribution of workers by NJCWM, by gender and match status at entry | | | | | | |
|---|---|---|---|---|---|---|
| | Overqualified | | | Well-matched | | |
| | Males | Females | Total | Males | Females | Total |
| 0 | 3,768 (79.7%) | 7,869 (84.7%) | 11,637 | 4,727 (66.6%) | 7,602 (68%) | 12,329 |
| 1 | 820 (17.3%) | 1,233 (13.3%) | 2,053 | 1,883 (26.6%) | 2,923 (26%) | 4,806 |
| 2 | 131 (2.8%) | 173 (1.9%) | 304 | 411 (5.8%) | 587 (5.3%) | 998 |
| 3 or + | 10 (0.2%) | 20 (0.1%) | 30 | 72 (1%) | 76 (0.7%) | 148 |
| Total | 4,729 | 9,295 | 14,024 | 7,093 | 11,188 | 18,281 |

Table 3.6 – Description of independent variables included in the model

| Variables | Description of variables |
|---|---|
| **Individual-level characteristics** | |
| Overqualified at entry (OQ) | = 1 if the worker is overqualified at entry, 0 otherwis |
| Female | = 1 for females, 0 otherwise |
| Age25-29 | = 1 if greater than 25 years old, 0 otherwise |
| Foreigner | = 1 for foreigners, 0 otherwise |
| Education Level: | |
| Bachelor | = 1 if the worker has a bachelor degree, 0 otherwise |
| Master | = 1 if the worker has a master degree, 0 otherwise |
| Field of study: | |
| Engineering and technology | = 1 if the field of study is engineering and technolog 0 otherwise |
| History, philosophy, etc. | = 1 if the field of study is History, philosophy etc., 0 otherwise |
| Arts | = 1 if the field of study is Arts, 0 otherwise |
| Business Science | = 1 if the field of study is Business, 0 otherwise |
| Education | = 1 if the field of study is Education, 0 otherwise |
| Journalism, media studies and communication | = 1 if the field of study is Journalism, 0 otherwise |
| Medicine | = 1 if the field of study is Medecine, 0 otherwise |
| Social Sciences | = 1 if the field of study is Social sciences, 0 otherwis |
| Cohort 2007 | = 1 if entered the labor market for the first time in 2007; 0 otherwise |
| **Job Characteristics** | |
| Fixed term contract | = 1 if fixed term contract, 0 otherwise |
| Wage Dispersion | Coef. of dispersion of real hourly earnings by occup. |
| Log real hourly earnings | Log of real hourly earnings in 2010 euros |
| Occupation categories 2-digit (ISCO08): | |
| Legislative power | = 1 if occup. is 11, 12, 13 or 14; 0 otherwise |
| Science and engineering | = 1 if occup. is 21, 22, 23, 24, 25 or 26; 0 otherwise |
| Technicians and associate professionals | = 1 if occup. is 31, 32, 33, 34, 35; 0 otherwise |
| Administrative staff | = 1 if occup. is 41, 42, 43 or 44; 0 otherwise |
| Service and sales | = 1 if occup. is 51, 52, 53 or 54; 0 otherwise |
| Skilled construction, industry sector | = 1 if occup. is 71, 72, 73, 74 or 75; 0 otherwise |
| Plant, machine operators, assemblers | = 1 if occup. is 81, 82 or 83; 0 otherwise |
| Unskilled workers | = 1 if occup. is 91, 92, 93, 94, 95 or 96; 0 otherwise |
| **Firm characteristics** | |
| Firm size | number of employees in the firm (in logs) |
| Urban location | = 1 if the firm is located in an urban area (districts of Porto or Lisbon); 0 otherwise |
| Industry dummies | Dummy variables for each 2-digit industry |

*Notes*: (i) all the independent variables are measured in the first job at entry; (ii) The hourly wages is regular payroll (base wage and regular payments) over normal hours worked in the reference month converted into 2010 constant prices using the Consumer Price Index (CPI).

Table 3.7 – Descriptive statistics, first job in 2006 or 2007

| Independent Variables | Mean | St. Dev. | Min. | Max |
|---|---|---|---|---|
| **Individual-level characteristics** | | | | |
| Overqualified at entry (OQ) | 0.434 | 0.495 | 0 | 1 |
| Female | 0.634 | 0.482 | 0 | 1 |
| Age25-29 (omitted category - age<25): | 0.650 | 0.477 | 0 | 1 |
| Foreigner | 0.019 | 0.138 | 0 | 1 |
| Education level (omitted category - bachelor): | | | | |
| Graduate | 0.844 | 0.362 | 0 | 1 |
| Master | 0.025 | 0.157 | 0 | 1 |
| Field of study: | | | | |
| Engineering and technology | 0.152 | 0.358 | | |
| History, philosophy, etc. | 0.035 | 0.183 | 0 | 1 |
| Arts | 0.021 | 0.143 | 0 | 1 |
| Business Science | 0.144 | 0.351 | 0 | 1 |
| Education | 0.070 | 0.256 | 0 | 1 |
| Journalism, media studies and communication | 0.022 | 0.145 | 0 | 1 |
| Medicine | 0.124 | 0.330 | 0 | 1 |
| Social sciences | 0.084 | 0.277 | 0 | 1 |
| Cohort 2007 | 0.515 | 0.500 | 0 | 1 |
| **Job Characteristics** | | | | |
| Fixed term contract | 0.672 | 0.469 | 0 | 1 |
| Wage Dispersion | 0.372 | 0.078 | 0.007 | 0.808 |
| Log real hourly earnings | 1.781 | 0.430 | 0.340 | 4.76 |
| Occupation categories 2-digit (ISCO-08): | | | | |
| Legislative power | 0.020 | 0.140 | 0 | 1 |
| Science and engineering | 0.544 | 0.498 | 0 | 1 |
| Technicians and associate professionals | 0.165 | 0.371 | 0 | 1 |
| Administrative staff | 0.193 | 0.395 | 0 | 1 |
| Service and sales | 0.058 | 0.233 | 0 | 1 |
| Skilled construction, industry sector | 0.008 | 0.089 | 0 | 1 |
| Plant, machine operators, assemblers | 0.003 | 0.057 | 0 | 1 |
| Unskilled workers | 0.009 | 0.092 | 0 | 1 |
| **Firm Characteristics** | | | | |
| Firm size (in logs) | 4.440 | 2.145 | 0 | 9.90 |
| Urban location | 0.767 | 0.423 | 0 | 1 |

Table 3.8 – PRM and OLS regressions

| Dependent variable: | NJC | | NJCWM | |
|---|---|---|---|---|
| | Poisson | OLS | Poisson | OLS |
| Overqualified ($OQ$) | −0.0245 | −0.012 | −0.4359*** | −0.1355*** |
| | (0.0213) | (0.0111) | (0.0311) | (0.0085) |
| Female | −0.0302* | −0.015* | −0.1398*** | −0.0447*** |
| | (0.0162) | (0.0086) | (0.0214) | (0.0068) |
| Age25-29 (omitted category: age<25) | 0.4143*** | 0.2321*** | 0.4262*** | 0.1350*** |
| | (0.0156) | (0.0092) | (0.0216) | (0.0073) |
| Foreigner | −0.3955*** | −0.1701*** | −0.5781*** | −0.1174*** |
| | (0.0678) | (0.0244) | (0.1132) | (0.0173) |
| Education level: (omitted category - bachelor): | | | | |
| Graduate | −0.0100 | −0.0065 | 0.0577* | 0.0187** |
| | (0.0213) | (0.0118) | (0.0316) | (0.0090) |
| Master | −0.4780*** | −0.1920*** | −0.3621** | −0.0833*** |
| | (0.0808) | (0.0258) | (0.1061) | (0.0218) |
| Field of study (omitted category - engineering & technology): | | | | |
| History, Philisophy, etc. | −0.0626 | −0.0312 | −0.3028*** | −0.0735*** |
| | (0.0423) | (0.0209) | (0.0681) | (0.0142) |
| Arts | −0.1805** | −0.0860** | −0.3516*** | −0.0885*** |
| | (0.05752) | (0.0254) | (0.0813) | (0.0185) |
| Business Science | 0.0880*** | 0.0477*** | −0.1488*** | −0.0457*** |
| | (0.0236) | (0.0132) | (0.0353) | (0.0097) |
| Education | -0.2120*** | −0.1020*** | −0.1152** | −0.0355** |
| | (0.0349) | (0.0157) | (0.0407) | (0.0135) |
| Journalism, media studies & communication | 0.0350 | 0.0221 | −0.0771 | −0.0248 |
| | (0.0509) | (0.0294) | (0.0757) | (0.0224) |
| Medicine | −0.0080 | −0.007 | 0.0646* | 0.0368** |
| | (0.0301) | (0.0159) | (0.0362) | (0.0135) |
| Social Sciences | −0.0441 | −0.0240 | −0.2326*** | −0.0677*** |
| | (0.0298) | (0.0152) | (0.0433) | (0.0114) |
| Cohort 2007 | −0.0683*** | −0.0370*** | −0.0723*** | −0.0229*** |
| | (0.0148) | (0.0080) | (0.0202) | (0.0063) |
| Fixed term contract | 0.2093*** | 0.1080*** | 0.1530*** | 0.0479*** |
| | (0.0179) | (0.0088) | (0.0231) | (0.0072) |
| Log real hourly earnings | −0.1028*** | −0.0518*** | 0.1862*** | 0.0559*** |
| | (0.0210) | (0.0105) | (0.0255) | (0.0085) |
| Wage Dispersion | 0.4776*** | 0.2665*** | −0.0637 | −0.0405 |
| | (0.1142) | (0.0600) | (0.1465) | (0.0479) |
| Firm size | −0.0114** | −0.0060** | −0.0421*** | −0.0124*** |
| | (0.0039) | (0.0021) | (0.0055) | (0.0075) |
| Urban Location | 0.0336* | 0.0171* | 0.0343 | 0.0101 |
| | (0.0182) | (0.0094) | (0.0238) | (0.0075) |
| $N$ | 32,305 | 32,305 | 32,305 | 32,305 |

*Notes*: (i) all regressions include a set of 2-dgit occupation dummies, non-defined fields of study and contract, as well as industry dummies; (ii) worker-cluster robust standard errors in parentheses; (iii) *, **, and *** denote significant at 10%, 5%, and 1% respectively; (iv) All independent variables are measured in the first job except wage dispersion at 3-digit.

Table 3.9 – APPENDIX A.5. ISCO-08 2-digit Occupation Classification

| ISCO-08 (2-digit) | Occupation categories |
|---|---|
| 11 | "Legislative power and executie bodies representatives, Seniors officials of Public Administration of special-interest org., entreprises directors, and managers" |
| 12 | "Administration and commercial directors" |
| 13 | "Production and specialised services directors" |
| 14 | "Hotels, food service, trade and others services directors" |
| 21 | "Physical sciences, mathematics, engineering and related techniques specialists" |
| 22 | "Health professionals" |
| 23 | "Teachers" |
| 24 | "Finance, accounting, administrative org., public and trade relations specialists" |
| 25 | "Information and communications technology specialists" |
| 26 | "Legal, social, artistic and cultural ,matters specialists" |
| 31 | "Science and engineering associate professionals" |
| 32 | "Health technicians ans associate professional" |
| 33 | "Financial, business and administration associated professionals" |
| 34 | "Legal, social, sport, cultural and related services, intermediate level technicians" |
| 35 | "Information and communications technicians" |
| 41 | "Office clerks, general secretaries and data keyboard clerks" |
| 42 | "Customer direct support staff" |
| 43 | "Data, accounting, statistical, financial services and material recording operators" |
| 44 | "Other clerical support workers" |
| 51 | "Personal service workers" |
| 52 | "Salespersons" |
| 53 | "Personal care and similar workers" |
| 54 | "Protective and safety services workers" |
| 61 | "Market-oriented farmers and skilled agricultural and farming of animals workers" |
| 62 | "Market-oriented skilled forestry, fishery and hunting workers" |
| 63 | "Subsistence farmers, fishers, hunters and gatherers" |
| 71 | "Building and related trades skilled workers, excluding electricians" |
| 72 | "Metal, machinery and related trades skilled workers" |
| 73 | "Printing, precision instruments manufacturing skilled workers jewelers, similar workers" |
| 74 | "Electrical and electronic trades skilled workers" |
| 75 | "Food processing, wood working, garment and other craft, related trade workers" |
| 81 | "Stationary plant and machine operators" |
| 82 | "Assemblers" |
| 83 | "Drivers and mobile plant operators" |
| 91 | "Cleaners and helpers" |
| 92 | "Agricultural, farming of animals, forestry and fishery not skilled workers" |
| 93 | "Minning, construction, manufacturing and transport not skilled workers" |
| 94 | "Food preparation assistants" |
| 95 | "Street vendors (excluding food), and street service workers" |
| 96 | "Refuse workers and other elementary workers" |

Table 3.10 – APPENDIX A.6. Classification of graduate and non graduate jobs by occupation at 2-digit (ISCO-08)

| ISCO-08 at 2-digit | NR-C.A | NR-C.I | RC | RM | NR-M.P | Job Classification | Mode of education (in years) |
|---|---|---|---|---|---|---|---|
| 11 | 1.6112 | 2.3021 | -0.8744 | -1.1852 | -1.3113 | Well-matched | 16 |
| 12 | 1.4161 | 1.6790 | -0.4775 | -1.2805 | -1.3568 | Well-matched | 16 |
| 13 | 1.1725 | 1.9956 | -0.5283 | -0.7799 | -0.8576 | Well-matched | 12 |
| 14 | 0.8870 | 2.1206 | -0.0007 | -0.6433 | -0.7125 | Well-matched | 12 |
| 21 | 1.8330 | 0.5355 | -0.2968 | -0.9816 | -0.9067 | Well-matched | 16 |
| 22 | 1.2768 | 1.0703 | -0.1268 | -0.8338 | -0.5323 | Well-matched | 16 |
| 23 | 1.1878 | 1.6510 | -1.1841 | -1.3690 | -1.2879 | Well-matched | 16 |
| 24 | 1.4964 | 0.8839 | -0.4062 | -1.4755 | -1.5128 | Well-matched | 16 |
| 25 | 1.5313 | -0.3552 | 0.8309 | -0.7611 | -0.8116 | Well-matched | 16 |
| 26 | 1.4444 | 0.3554 | -0.3664 | -1.0821 | -1.1222 | Well-matched | 16 |
| 31 | 0.4181 | -0.0219 | 0.7685 | 0.6341 | 0.5347 | Overqualified | 12 |
| 32 | 0.6943 | 0.4512 | 0.7116 | -0.0649 | -0.0281 | Overqualified | 12 |
| 33 | 0.5005 | 0.1811 | 0.6003 | -0.7882 | -0.8097 | Overqualified | 12 |
| 34 | 0.1737 | 0.3917 | -0.2179 | -0.4523 | -0.2420 | Overqualified | 12 |
| 35 | 0.7429 | -0.7542 | 0.5334 | -0.1180 | -0.2961 | Overqualified | 12 |
| 41 | -0.4103 | -1.1258 | 1.1156 | -0.5818 | -1.0274 | Overqualified | 12 |
| 42 | -0.3134 | -0.2444 | 1.5457 | -0.4210 | -0.9211 | Overqualified | 12 |
| 43 | 0.0011 | -0.5628 | 1.5947 | -0.2388 | -0.7711 | Overqualified | 12 |
| 44 | -0.4875 | -0.9898 | 1.3798 | -0.3124 | -0.6674 | Overqualified | 12 |
| 51 | -0.6836 | 0.2696 | -0.0452 | -0.1230 | -0.1286 | Overqualified | 9 |
| 52 | -0.8447 | -0.0465 | -0.7103 | -0.5301 | -0.6213 | Overqualified | 9 |
| 53 | 0.0656 | 0.5126 | -0.2518 | -0.7084 | -0.5967 | Overqualified | 9 |
| 54 | 0.4098 | 1.0405 | 0.8002 | -0.2754 | 0.6827 | Well-matched | 9 |
| 61 | -0.9095 | -0.7624 | -1.0813 | 0.5696 | 1.1153 | Overqualified | 4 |
| 62 | -0.7173 | -0.3877 | -1.1520 | 0.6063 | 1.2187 | Overqualified | 4 |
| 71 | -0.2670 | -0.0693 | -0.0957 | 0.7846 | 1.3045 | Overqualified | 4 |
| 72 | -0.4406 | -0.6071 | 0.4094 | 1.3140 | 1.2720 | Overqualified | 6 |
| 73 | -0.4484 | -1.2835 | 0.6363 | 1.0716 | 0.5413 | Overqualified | 6 |
| 74 | 0.2938 | 0.1091 | 0.1995 | 0.3035 | 1.2191 | Overqualified | 9 |
| 75 | -0.9220 | -1.0798 | 0.5329 | 1.2715 | 0.7187 | Overqualified | 4 |
| 81 | -0.7998 | -0.7690 | 0.7995 | 2.1755 | 0.9126 | Overqualified | 6 |
| 82 | -0.5707 | -0.7698 | 0.4407 | 1.7236 | 1.121 | Overqualified | 9 |
| 83 | -0.5025 | -0.2652 | 0.3681 | 1.1175 | 1.4862 | Overqualified | 4 |
| 91 | -1.5522 | -1.5267 | -0.4165 | 0.8440 | 0.7375 | Overqualified | 4 |
| 92 | -1.0138 | -0.7880 | -0.9747 | 0.9674 | 1.1311 | Overqualified | 4 |
| 93 | -0.6421 | -0.7176 | 0.3231 | 1.2764 | 1.3710 | Overqualified | 4 |
| 94 | -1.4337 | -0.7091 | 0.2734 | 0.7490 | 0.2171 | Overqualified | 4 |
| 95 | -1.5927 | 0.6046 | -4.0765 | -2.1243 | -1.5618 | Overqualified | 9 |
| 96 | -1.1992 | -1.1628 | 0.6294 | 0.8938 | 1.3202 | Overqualified | 9 |

*Note*: the scores are standardized to mean 0 and standard deviation 1.